(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 April 2003 (17.04.2003)

PCT

(10) International Publication Number
WO 03/031932 A2

(51) International Patent Classification[7]: G01N

(21) International Application Number: PCT/US02/31988

(22) International Filing Date: 7 October 2002 (07.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/327,537 5 October 2001 (05.10.2001) US

(71) Applicant *(for all designated States except US)*: **CASE WESTERN RESERVE UNIVERSITY** [US/US]; 10900 Euclid Avenue, Cleveland, OH 44106-4971 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **MARKOWITZ, Sanford, D.** [US/US]; 28950 Bolingbrook Road, Pepper Pike, OH 44124 (US).

(74) Agents: **VARMA, Anita** et al.; Ropes & Gray, One International Place, Boston, MA 02110-2624 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/031932 A2

(54) Title: METHODS AND COMPOSITIONS FOR DETECTING COLON CANCERS

(57) Abstract: This application describes methods and compositions for detecting and treating HLTF-associated neoplasia. Differential mehylation of the HLTF nucleotide sequences has been observed in HLTF-associated neoplasia such as colon neoplasia.

METHODS AND COMPOSITIONS FOR DETECTING COLON CANCERS

Cross-Reference to Related Applications

This application claims the benefit of priority of U.S. Provisional Application No. 60/327,537, filed October 5, 2001, the specification of which is incorporated by reference herein in its entirety.

Background

In 2001, over 1.2 million new cases of human cancer will be diagnosed and over 0.5 million people will die from cancer (American Cancer Society estimate). Despite this, more people than ever are living with and surviving cancer. In 1997, for example, approximately 8.9 million living Americans had a history of cancer (National Cancer Institute estimate). People are more likely to survive cancer if the disease is diagnosed at an early stage of development, since treatment at that time is more likely to be successful. Early detection depends upon availability of high-quality methods. Such methods are also useful for determining patient prognosis, selecting therapy, monitoring response to therapy and selecting patients for additional therapy. Consequently, there is a need for cancer diagnostic methods that are specific, accurate, minimally invasive, technically simple and inexpensive.

Colorectal cancer (i.e., cancer of the colon or rectum) is one particularly important type of human cancer. Colorectal cancer is the second most common cause of cancer mortality in adult Americans (Landis, et al., 1999, CA Cancer J Clin, 49:8-31). Approximately 40% of individuals with colorectal cancer die. In 2001, it is estimated that there will be 135,400 new cases of colorectal cancer (98,200 cases of colon and 37,200 cases of rectal cancer) and 56,700 deaths (48,000 colon cancer and 8,800 rectal cancer deaths) from the disease (American Cancer Society). As with other cancers, these rates can be decreased by improved methods for diagnosis.

Although methods for detecting colon cancer exist, the methods are not ideal. Digital rectal exams (i.e., manual probing of rectum by a physician), for example, although relatively inexpensive, are unpleasant and can be inaccurate. Fecal occult blood testing (i.e., detection of blood in stool) is nonspecific because blood in the stool has multiple causes. Colonoscopy and sigmoidoscopy (i.e., direct examination

of the colon with a flexible viewing instrument) are both uncomfortable for the patient and expensive. Double-contrast barium enema (i.e., taking X-rays of barium-filled colon) is also an expensive procedure, usually performed by a radiologist.

Because of the disadvantages of existing methods for detecting colon cancer, new methods are needed.

Summary Of The Invention

In accordance with the present invention, new methods are provided for detection of colon cancer. In one aspect, the method comprises assaying for the presence of differentially methylated HLTF nucleotide sequences in a tissue sample or a bodily fluid sample from a subject. Preferred bodily fluids include blood, serum, plasma, a blood-derived fraction, stool, colonic effluent or urine. In one embodiment, the method involves restriction enzyme/methylation-sensitive PCR. In another embodiment, the method comprises reacting DNA from the sample with a chemical compound that converts non-methylated cytosine bases (also called "conversion-sensitive" cytosines), but not methylated cytosine bases, to a different nucleotide base. In a preferred embodiment, the chemical compound is sodium bisulfite, which converts unmethylated cytosine bases to uracil. The compound-converted DNA is then amplified using a methylation-sensitive polymerase chain reaction (MSP) employing primers that amplify the compound-converted DNA template if cytosine bases within CpG dinucleotides of the DNA from the sample are methylated. Production of a PCR product indicates that the subject has colon cancer. Other methods for assaying for the presence of methylated DNA are known in the art.

In another aspect, the method comprises assaying for decreased levels of an HLTF transcript in the sample. Examples of such assays include RT-PCR assays which employ primers that derived from the coding sequence of HLTF.

In another aspect, the present invention provides a detection method for prognosis of a colon cancer in a subject known to have or suspected of having colon

cancer. Such method comprises assaying for the presence of differentially methylated HLTF nucleotide sequences in a tissue sample or bodily fluid from the subject. In certain cases, presence of differentially methylated HLTF nucleotide sequences in the tissue sample bodily fluid indicates that the subject is a good candidate for a particular therapy. In other cases, presence of the differentially methylated HLTF nucleotide sequences in the tissue sample or bodily fluid indicates that the colon cancer has a poor prognosis or the subject is a candidate for more aggressive therapy.

In another aspect, the present invention provides a method for monitoring over time the status of colon cancer in a subject. The method comprises assaying for the presence of differentially methylated HLTF nucleotide sequences in a tissue sample or bodily fluid taken from the subject at a first time and in a corresponding bodily fluid taken from the subject at a second time. Absence of differentially methylated HLTF nucleotide sequences from the bodily fluid taken at the first time and presence of differentially methylated HLTF nucleotide sequences in the bodily fluid taken at the second time indicates that the cancer is progressing. Presence of differentially methylated HLTF nucleotide sequences in the tissue sample or bodily fluid taken at the first time and absence of differentially methylated HLTF nucleotide sequences from the tissue sample or bodily fluid taken at the second time indicates that the cancer is regressing.

In another aspect the present invention provides a method for evaluating therapy in a subject suspected of having or having colon cancer. The method comprises assaying for the presence of methylated HLTF promoter DNA in a tissue sample or bodily fluid taken from the subject prior to therapy and a corresponding bodily fluid taken from the subject during or following therapy. Loss of methylated HLTF promoter DNA or a decrease in methylation of HLTF promoter DNA in the sample taken after or during therapy as compared to the levels or HLTF promoter DNA in the sample taken before therapy is indicative of a positive effect of the therapy on cancer regression in the treated subject

The present invention also provides nucleotide primer sequences for use in the methylation-sensitive PCR assay.

The present invention also provides a method of inhibiting or reducing growth of colon cancer cells. The method comprises increasing the levels of the protein encoded by HLTF in colon cancer cells. In one embodiment, the cells are contacted with the HLTF protein or a biologically active equivalent or fragment thereof under conditions permitting uptake of the protein or fragment. In another embodiment, the cells are contacted with a nucleic acid encoding the HLTF protein and (ii) a promoter active in the cancer cell, wherein the promoter is operably linked to the region encoding the HLTF protein, under conditions permitting the uptake of the nucleic acid by the cancer cell. In another embodiment, the method comprises demethylating the methylated HLTF promoter DNA.

In one aspect, the application provides isolated or recombinant HLTF nucleotide sequences that are at least 80%, 85%, 90%, 95%, 98%, 99% or identical to the nucleotide sequence of any one of SEQ ID NOs: 2-4 and 21, fragments of said sequences that are 10, 15, 20, 25, 50, 100, or 150 base pairs in length wherein the HLTF nucleotide sequences are differentially methylated in an HLTF-associated disease cell.

Brief Description Of The Drawings

Figure 1 illustrates HLTF silencing in colon cancer cell lines. **A.** HLTF RNA expression. Shown is an RT-PCR assay for HLTF expression in colon cancer cell lines. **B.** HLTF expression reactivation. Shown is an RT-PCR assay for HLTF expression in colon cancer cell lines treated (+) or untreated (-) with 5-azacytdine (5-azaC). Cell lines V429 and V503 are controls with constitutive HLTF expression. 5-azaC treatment reactivates HLTF expression in cell lines RCA, V457, SW480, V5, V6, and V432.

Figure 2 shows methylation of HLTF 5' genomic sequence. **A.** Diagram of the HLTF 5' genomic region. CpG sites are shown with circles. Shaded circles represent CpG sites that are tested in MS-PCR assays. Hatched circles represent

CpG sites that overlap HpaII restriction sites. **B.** MS-PCR assay of the HLTF 5' genomic sequence. Shown are the results of MS-PCR assay of the HLTF 5' genomic sequence by using primers specific for amplification of either methylated (M) or unmethylated (U) templates. **C.** HLTF MS-PCR of matched cell lines and tissues. Shown are the results of HLTF MS-PCR assay of colon cancer cell lines (C), matched antecedent tumor tissue (T) or matched normal colon mucosa (N).

Figure 3 shows methylation status of the HLTF 5' genomic region in primary tumors and matched normal tissues. Shown are the results of MS-PCR assay of the HLTF 5' genomic region in matched paired tumor (T) and normal (N) colon tissues samples amplified with primers specific for methylated (M) or unmethylated (U) templates.

Figure 4 shows correlation of HLTF 5' genomic region methylation with age. **A.** In colon cancer tumors and cell lines, . **B.** In normal colon tissues.

Figure 5 shows correlation of HLTF 5' genomic region methylation with tumor site (**A, B**) or with tumor stage (**C, D**). Shown in **A** and **B** are percentage (%) of colon neoplasms (tumors and cell lines) in each category defined by location of the tumor in the colon and HLTF methylation status. Shown in **C** and **D** are percentage (%) of colon neoplasms (tumors and cell lines) in each category defined by clinical stage of the colon tumor and HLTF methylation status.

Figure 6 illustrates correlation of HLTF methylation with the CpG island methylator phenotype (i.e., CIMP status) and with hMLH1 methylation. **A.** Shown are the numbers of primary colon cancers in each of the categories defined by combined HLTF methylation and CIMP status. **B.** Shown are the numbers of colon cancers (tumors and cell lines) in each category defined by combined hMLH1 and HLTF methylation status.

Figure 7 shows HLTF colony suppressor activity. Shown are the number of G418 resistant colonies arising from transfection with an HLTF expression vector (HLTF) or a control empty expression vector (pcDNA) in HLTF unmethylated and expressing FET, V364 and V429 cells (**B**) as compared to HLTF methylated and

deficient V457, V8-2, and RCA cells (A). C. Anti-V5 western blot assay of V5-epitope tagged HLTF introduced by transient transfection into HLTF methylated versus unmethylated cells. Control cells were transfected with an empty expression vector (pcDNA3.1).

Figure 8 shows a diagram of the HLTF 5' genomic region. CpG sites are shown with circles and stems. Hatched circles represent CpG sites that overlap HpaII restriction sites. Sequences that are complementary to PCR primers that were used to selectively amplify the methylated but not unmethylated HLTF 5' genomic sequence after digestion with HpaII are designated by the location of the arrows corresponding to forward PCR primer 1277F and reverse PCR primer 1724R. Shaded circles represent the CpG sites that are tested by MS-PCR assay primers described as examples in this application. Locations of specific primers used in the specific MS-PCR assays described as examples in this application are indicated with arrows and correspond to forward PCR primer 1352MF, designed as a forward primer for amplification of bisulfite converted sense sequences of duplex DNA derived from the methylated parental sense strand, and reverse primers 1606MR and 1627MR, designed as reverse primers for amplification of bisulfite converted sequences of duplex DNA derived from methylated parental sense strand. Primers 1352MF(ASS) and 1607MR(ASS) indicate the forward and reverse primers for amplification of duplex DNA derived from bisulfite converted sequences of methylated parental antisense strands. Further, the control primers (indicated as UF or UR) that are used to detect the unmethylated HLTF template in an MSP are also indicated in the diagram.

Figure 9 shows the structure of the 5' region of HLTF at 3 levels of resolution. The top panel (**A**) depicts residues 0-4500 spanning exons 1 and exons 2, and depicting the position of three Alu repeats in which Alu1 is upstream of exon 1, Alu 2 is in intron 1, and Alu 3 is in intron 3. Balloons designate the positions of CpG dinucleotide sequences. The second panel (**B**) depicts at higher magnification the structure between residues 0-3000, spanning Alu1, exon1, and Alu2. The third panel (**C**) shows at higher magnification the region from residues 550 to 2459 that is between Alu1 and Alu2.

Figure 10 summarizes the results of the sequencing across HLTF genomic residues 0-3000 of bisulfite converted genomic DNA from 6 different samples: Vaco5 (an HLTF silenced colon cancer cell line), Vaco206 (an HLTF expressing colon cancer cell line), and 4 normal colon epithelial tissue samples (19-11N, 587N, 421N, and 406N). Multiple individual DNA clones were sequenced for each of the bisulfite converted samples. The positions of Alu1, HLTF exon1, and Alu2 are shown in boxes. Open balloons denote CpG residues where cytosine methylation was found to be constitutive in normal colonic tissue. In both Alu1 and Alu2, all CpG residues are constitutively methylated in normal tissue. Additionally, CpG residues that are 3' of Alu1 from bases 550-1200 were also all found to be constitutively methylated. A differentially methylated region, that is methylated in HLTF silenced Vaco5 and is, in general, not methylated in normal colon or in HLTF expressing cancer is defined by the CpG dinucleotides lying between residues 1200 and 2600. The differential methylation of 5 HpaII sites between residues 1277 and 1742 was independently confirmed in multiple additional normal and cancer samples by assays of the resistance of these sites to HpaII digestion. Additionally, 3 sets of MS-PCR primers were designed to assay the methylation status of residues between 1352 and 1672 (as shown on the figure), and these MS-PCR assays also confirmed that these residues were unmethylated in normal colon tissue and in HLTF expressing colon cancers, but were methylated in HLTF silenced colon cancers.

Figure 11 shows a final diagramatic summary of the structure of the region between Alu1 and Alu2 repeats, and designates the boundaries of the base pair 1200-2500 region outside of which there is constitutive methylation of the adjacent CpG dinucleotides.

Figure 12 summarizes for HLTF gene residues 550 to 2500 the primary results of the bisulfite sequencing in the 6 different samples (corresponding to the region flanked by the Alu1 and Alu2 repeats). The balloons indicate the position of the CpG dinucleotides. Data from the 6 samples is summarized by the 6 lines. At each CpG residue, an open circle indicates the residue was unmethylated in that sample, a black filled circle indicates the residue was methylated in every bisulfite converted clone derived from that sample, and a grey filled circle indicates that the residue was

methylated in from 20%-70% of the clones derived from that sample. In the HLTF silenced Vaco5 sample, there basically is methylation of every residue from base pair 550 to 2500. However, the residues from base pairs 550 to 1200 are also seen to be methylated in the normal colon samples as well as in the HLTF expressing cancer cell line Vaco206. The residues between 1200 and 2500 essentially define a region that is differentially methylated in the HLTF silenced Vaco5 sample. Note is made of slight methylation of residues 1200 to 1400 in one normal sample, and of residues 2300 to 2500 in Vaco206. However, this slight methylation is distinguished from the core methylation of HLTF silenced cancers by all of the MS-PCR assays described in this application as well as by the described assay of methylation of the HpaII sites that span the 1277 to 1742 interval.

Figure 13 shows the results of bisulfite sequencing of multiple individually derived DNA clones derived from bisulfite treated DNA from the Vaco5 colon cancer cell lines which do not express HLTF. In each of these Vaco5 derived DNA clones, essentially all of the CpG sites are methylated in the region from residues 550 to 2500.

Figure 14 shows the amino acid sequence (SEQ ID NO: 1) of human HLTF protein.

Figure 15 shows the 5' genomic sequence of human HLTF gene (residues 1 to 3000, sense strand, SEQ ID NO: 2). The underlined region (residues 1250-1800, SEQ ID NO: 4) was tested by methylation specific PCR and by sensitivity to HpaII digestion. Alu1 and Alu2 regions are in bold. The start ATG is underlined and in bold, with the A at position 1757. There is currently one complete GeneBank entry "AC021059: Homo sapiens 3 BAC RP11-464E15 (Roswell Park Cancer Institute Human BAC Library) complete sequence", which contains the human HLTF gene. Residues 1-3000 correspond to positions 119396-116395 of this genomic clone (AC021059), and the A of the ATG relative to AC020159 will be 117640.

Figure 16 shows the sequence of residue 600-2600 (SEQ ID NO: 3) that includes the differentially methylated region of residues 1200-2600, and a portion of the non-Alu constitutively methylated region within residues 600-1200.

Figure 17 shows the HLTF nucleotide sequence base pairs 1250-1800 (SEQ ID NO: 4). CpG dinucleotides that are sites of cytosine methylation are in bold. Those CpG sites that are HpaII restriction sites are shown in italic and underlined. Sequences that are complementary to PCR primers that were used to selectively amplify the methylated but not unmethylated HLTF DNA after digestion with HpaII are shown as bold arrows. Those CpG site that are tested by specific MS-PCR assays described as examples in this invention are underlined. Parent sequences that were used to design specific MS-PCR primers that amplified methylated but not unmethylated templates following conversion with bisulfite are shown in smaller arrows.

Figure 18 shows the sequences following bisulfite conversion of DNA derived from the sense strand of methylated template (top panel) and unmethylated template (bottom panel) of the HLTF 5' genomic sequence residues 1250-1800. CpG dinucleotides that are sites of cytosine methylation are in bold. Those CpG sites that are HpaII restriction sites are shown in italic and underlined. Sequences that are complementary to PCR primers that were used to selectively amplify the methylated but not unmethylated HLTF DNA after digestion with HpaII are shown as bold arrows. Those CpG site that are tested by specific MS-PCR assays described as examples in this invention are shown underlined. Sequences that were used to design specific MS-PCR primers that amplified methylated but not unmethylated templates following conversion with bisulfite are shown in smaller arrows.

Figure 19 shows the corresponding complementary strands of the bisulfite-converted HLTF DNA base pairs 1250-1800 (methylated and unmethylated templates, as shown in Figure 18). CpG dinucleotides that are sites of cytosine methylation are in bold. Those CpG sites that are HpaII restriction sites are shown in italic and underlined. Sequences that are complementary to PCR primers that were used to selectively amplify the methylated but not unmethylated HLTF DNA after digestion with HpaII are shown as bold arrows. Those CpG site that are tested by specific MS-PCR assays described as examples in this invention are shown underlined. Sequences that were used to design specific MS-PCR primers that

amplified methylated but not unmethylated templates following conversion with bisulfite are shown in smaller arrows.

Figure 20 shows primer sequences for amplifying HLTF. **A.** Forward PCR primer 1277F (SEQ ID NO: 9) and reverse PCR primer 1724R (SEQ ID NO: 10) selectively amplify the methylated but not unmethylated HLTF sequence after digestion with HpaII. Unmethylated DNAs are cut by HpaII and so cannot be PCR amplified **B** and **C** show primer sets for amplifying bisulfite-converted sense sequences of the duplex methylated HLTF DNA: forward PCR primer 1352MF (SEQ ID NO: 11) and reverse primer 1606MR (SEQ ID NO: 12); forward PCR primer 1352MF (SEQ ID NO: 11) and reverse primer 1627MR (SEQ ID NO: 15). **D** shows primer sets for amplifying bisulfite-converted antisense sequences of the duplex methylated HLTF DNA: forward primer 1352MF(ASS) (SEQ ID NO: 17) and reverse primer 1607MR(ASS) (SEQ ID NO: 18). Sequences underlined in **B, C,** and **D** are the control primer sets used to amply bisulfite-converted sequences (sense or antisense) of the duplex unmethylated HLTF DNA (designated as UF or UR): forward PCR primer 1347UF (SEQ ID NO: 13) and reverse primer 1610UR (SEQ ID NO: 14); forward PCR primer 1347UF (SEQ ID NO: 13) and reverse primer 1631UR (SEQ ID NO: 16); forward primer 1349UF(ASS) (SEQ ID NO: 19) and reverse primer 1611UR(ASS) (SEQ ID NO: 20).

Figure 21 shows the HLTF 5' genomic sequence (residues 1200-2500, sense strand, SEQ ID NO: 21). The region is differentially methylated as shown in figure 10.

Figure 22 shows the sequence (SEQ ID NO: 22), derived from bisulfite conversion of DNA derived from the sense strand of methylated template of the HLTF 5' genomic sequence residues 1200-2500 (i.e., SEQ ID NO: 21). The underlined region (residues 1250-1800) was tested by MSP assay. The start ATG is in bold.

Figure 23 shows the sequence (SEQ ID NO: 23), derived from bisulfite conversion of DNA derived from the sense strand of unmethylated template of the HLTF 5' genomic sequence residues 1200-2500 (i.e., SEQ ID NO: 21). The underlined region (residues 1250-1800) was tested by MSP assay. The start ATG is in bold.

Figure 24 shows the sequence (SEQ ID NO: 24), derived from bisulfite conversion of DNA derived from the antisense strand of methylated template of the HLTF 5' genomic sequence residues 1200-2500 (i.e., SEQ ID NO: 21). The underlined region (residues 1250-1800) was tested by MSP assay. The start ATG is in bold.

Figure 25 shows the sequence (SEQ ID NO: 25), derived from bisulfite conversion of DNA derived from the antisense strand of unmethylated template of the HLTF 5' genomic sequence residues 1200-2500 (i.e., SEQ ID NO: 21). The underlined region (residues 1250-1800) was tested by MSP assay. The start ATG is in bold.

Figure 26 shows digramatically the positions of newly designed primer sets 4-7 for detecting differential methylation of the 3' HLTF region that is beyond the previously-tested region.

Figure 27 shows the sequences of the proposed primer sets 4-8. MSP4, MSP5, MSP7, and MSP8 are primer sets for amplifying bisulfite-converted antisense sequences of the duplex methylated HLTF DNA, including: forward primer 1P-HLTF1581MF(ASS) (SEQ ID NO: 26) and reverse primer 13P-HLTF1713MR(ASS) (SEQ ID NO: 27); forward primer 1P-HLTF1581MF(ASS) (SEQ ID NO: 26) and reverse primer 5P-HLTF1827MR(ASS) (SEQ ID NO: 30); forward primer 9P-HLTF1893MF(ASS) (SEQ ID NO: 36) and reverse primer ALU(MB)2133FR(ASS) (SEQ ID NO: 37); forward primer 15P-HLTF2201MF(ASS) (SEQ ID NO: 40) and reverse primer 11P-HLTF2400MR(ASS) (SEQ ID NO: 41). MSP6 are primer sets for amplifying bisulfite-converted sense sequences of the duplex methylated HLTF DNA, including forward primer 3P-HLTF1621MF (SEQ ID NO: 32) and reverse primer 7P-HLTF1873MR (SEQ ID NO: 33). Sequences underlined are the control primer sets used to amplify bisulfite-converted sequences (sense or antisense) of the duplex unmethylated HLTF DNA (designated as UF or UR), including: forward primer 2P-HLTF1575UF(ASS) (SEQ ID NO: 28) and reverse primer 14P-HLTF1728UR(ASS) (SEQ ID NO: 29); forward primer 2P-HLTF1575UF(ASS) (SEQ ID NO: 28) and reverse primer 6P-HLTF1829UR(ASS) (SEQ ID NO: 31); forward primer 4P-HLTF1614UF (SEQ ID NO: 34) and reverse primer 8P-HLTF1878UR (SEQ ID

NO: 35); forward primer 10P-HLTF1890UF(ASS) (SEQ ID NO: 38) and reverse primer ALU(MB)2133FR(ASS) (SEQ ID NO: 37); forward primer 16P-HLTF2197UF(ASS) (SEQ ID NO: 42) and reverse primer 12P-HLTF2403UR(ASS) (SEQ ID NO: 43).

Figure 28 shows a region of the Genomic clone AC021059 (residues 58381-120901) (SEQ ID NO: 39), encompassing the HLTF gene. The HLTF gene is located on the antisense strand of the clone.

Detailed Description Of The Invention

I. Definitions

For convenience, certain terms employed in the specification, examples, and appended claims are collected here. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

The articles "a" and "an" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

The terms "adenoma", "colon adenoma," and "polyp" are used herein to describe any precancerous neoplasia of the colon.

The term "colon" as used herein is intended to encompass the right colon (including the cecum), the transverse colon, the left colon, and the rectum.

The terms "colorectal cancer" and "colon cancer" are used interchangeably herein to refer to any cancerous neoplasia of the colon (including the rectum, as defined above).

The term "blood-derived fraction" herein refers to a component or components of whole blood. Whole blood comprises a liquid portion (i.e., plasma) and a solid portion (i.e., blood cells). The liquid and solid portions of blood are each comprised of multiple components; e.g., different proteins in plasma or different cell

types in the solid portion. One of these components or a mixture of any of these components is a blood-derived fraction as long as such fraction is missing one or more components found in whole blood.

"Cells," "host cells" or "recombinant host cells" are terms used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

The terms "compound", "test compound," "agent", and "molecule" are used herein interchangeably and are meant to include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, natural product extract libraries, and any other molecules (including, but not limited to, chemicals, metals, and organometallic compounds).

The term "compound-converted DNA" herein refers to DNA that has been treated or reacted with a chemical compound that converts unmethylated C bases in DNA to a different nucleotide base. For example, one such compound is sodium bisulfite, which converts unmethylated C to U. If DNA that contains conversion-sensitive cytosine is treated with sodium bisulfite, the compound-converted DNA will contain U in place of C. If the DNA which is treated with sodium bisulfite contains only methylcytosine, the compound-converted DNA will not contain uracil in place of the methylcytosine.

The term "de-methylating agent" as used herein refers agents that restore activity and/or gene expression of target genes silenced by methylation upon treatment with the agent Examples of such agents include without limitation 5-azacytidine, 5-aza-2'-deoxycytidine,

As used herein, the phrase "gene expression" or "protein expression" includes any information pertaining to the amount of gene transcript or protein present in a sample, as well as information about the rate at which genes or proteins

are produced or are accumulating or being degraded (e.g., reporter gene data, data from nuclear runoff experiments, pulse-chase data etc.). Certain kinds of data might be viewed as relating to both gene and protein expression. For example, protein levels in a cell are reflective of the level of protein as well as the level of transcription, and such data is intended to be included by the phrase "gene or protein expression information." Such information may be given in the form of amounts per cell, amounts relative to a control gene or protein, in unitless measures, etc.; the term "information" is not to be limited to any particular means of representation and is intended to mean any representation that provides relevant information. The term "expression levels" refers to a quantity reflected in or derivable from the gene or protein expression data, whether the data is directed to gene transcript accumulation or protein accumulation or protein synthesis rates, etc.

The term "detection" is used herein to refer to any process of observing a marker, or a change in a marker (such as for example the change in the methylation state of the marker), in a biological sample, whether or not the marker or the change in the ,marker is actually detected. In other words, the act of probing a sample for a marker or a change in the marker, is a "detection" even if the marker is determined to be not present or below the level of sensitivity. Detection may be a quantitative, semi-quantitative or non-quantitative observation.

"differentially methylated HLTF nucleotide sequence" refers to a region of the HLTF nucleotide sequence that is found to be methylated in an HLTF-associated neoplasia such as a region of the HLTF nucleotide sequence that is found to be methylated in colon cancer tissues or cell lines, but not methylated in the normal tissues or cell lines. For example, Figure 10 delineates certain HLTF regions that are differentially methylated regions. Illustrative examples of such differentially methylated HLTF regions are set forth in SEQ ID NOs: 4 and 21.

"Expression vector" refers to a replicable DNA construct used to express DNA which encodes the desired protein and which includes a transcriptional unit comprising an assembly of (1) genetic element(s) having a regulatory role in gene expression, for example, promoters, operators, or enhancers, operatively linked to

(2) a DNA sequence encoding a desired protein (in this case, an HLTF protein) which is transcribed into mRNA and translated into protein, and (3) appropriate transcription and translation initiation and termination sequences. The choice of promoter and other regulatory elements generally varies according to the intended host cell. In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. In the present specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

In the expression vectors, regulatory elements controlling transcription or translation can be generally derived from mammalian, microbial, viral or insect genes. The ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants may additionally be incorporated. Vectors derived from viruses, such as retroviruses, adenoviruses, and the like, may be employed.

The terms "healthy", "normal," and "non-neoplastic" are used interchangeably herein to refer to a subject or particular cell or tissue that is devoid (at least to the limit of detection) of a disease condition, such as a neoplasia, that is associated with HLTF such as for example neoplsia associated with silencing of HLTF gene expression due to methylation. These terms are often used herein in reference to tissues and cells of the colon. Thus, for the purposes of this application, a patient with severe heart disease but lacking a HLTF silencing-associated disease would be termed "healthy."

"HLTF-associated neoplasia" refers to neoplasia associated with reduced expression or no expression of the HLTF gene. Examples of HLTF-associated neoplasia include gastro-intestinal neoplasia , colon neoplasia etc.

"HLTF-associated proliferative disorder" refers to a disease that is associated with either reduced expression or over-expression of the HLTF gene.

"HLTF-methylation target regions" as used herein refer to those regions of HLTF that are found to be methylated. These regions include nucleotide regions that may be either constitutively or differentially methylated regions. For example, Figure 10 discloses an HLTF region wherein certain regions of the sequence are constitutively methylated and certain other regions are differentially methylated regions. Illustrative examples of such HLTF methylation target regions are set forth in SEQ ID NO: 2-3 and 39.

"HLTF-nucleotide sequence" or "HLTF-nucleic acid sequence" as used herein refers to the HLTF-genomic sequences as set forth in SEQ ID NO: 39 and to the 5'-genomic flanking regulatory regions as set forth in SED ID NOs: 2-4 and 21.

"HLTF-silencing associated diseases" as used herein includes HLTF-associated neoplasia.

"Homology" or "identity" or "similarity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology and identity can each be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When an equivalent position in the compared sequences is occupied by the same base or amino acid, then the molecules are identical at that position; when the equivalent site occupied by the same or a similar amino acid residue (e.g., similar in steric and/or electronic nature), then the molecules can be referred to as homologous (similar) at that position. Expression as a percentage of homology/similarity or identity refers to a function of the number of identical or similar amino acids at positions shared by the compared sequences. A sequence which is "unrelated" or "non-homologous" shares less than 40% identity, preferably less than 25% identity with a sequence of the present invention. In comparing two sequences, the absence of residues (amino acids or nucleic acids) or presence of extra residues also decreases the identity and homology/similarity.

The term "homology" describes a mathematically based comparison of sequence similarities which is used to identify genes or proteins with similar functions or motifs. The nucleic acid and protein sequences of the present invention may be used as a "query sequence" to perform a search against public databases to,

for example, identify other family members, related sequences or homologs. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, et al. (1990) J Mol. Biol. 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score=100, wordlength=12 to obtain nucleotide sequences homologous to nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score=50, wordlength=3 to obtain amino acid sequences homologous to protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., (1997) Nucleic Acids Res. 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and BLAST) can be used. See http://www.ncbi.nlm.nih.gov.

As used herein, "identity" means the percentage of identical nucleotide or amino acid residues at corresponding positions in two or more sequences when the sequences are aligned to maximize sequence matching, i.e., taking into account gaps and insertions. Identity can be readily calculated by known methods, including but not limited to those described in (Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48: 1073, 1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available computer programs. Computer program methods to determine identity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., Nucleic Acids Research 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S. F. et al., J. Molec. Biol. 215: 403-410 (1990) and Altschul et al. Nuc. Acids Res. 25: 3389-3402 (1997)). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, Md. 20894;

Altschul, S., et al., J. Mol. Biol. 215: 403-410 (1990)). The well known Smith Waterman algorithm may also be used to determine identity.

The term "including" is used herein to mean, and is used interchangeably with, the phrase "including but not limited to."

The term "isolated" as used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules in a form which does not occur in nature. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state.

The term "methylation-sensitive PCR" (i.e., MSP) herein refers to a polymerase chain reaction in which amplification of the compound-converted template sequence is performed. Two sets of primers are designed for use in MSP. Each set of primers comprises a forward primer and a reverse primer. One set of primers, called methylation-specific primers (see below), will amplify the compound-converted template sequence if C bases in CpG dinucleotides within the HLTF DNA are methylated. Another set of primers, called unmethylation-specific primers (see below), will amplify the compound-converted template sequences if C bases in CpG dinucleotides within the HLTF DNA are not methylated.

As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single-stranded (such as sense or antisense) and double-stranded polynucleotides.

"Operably linked" when describing the relationship between two DNA regions simply means that they are functionally related to each other. For example, a promoter or other transcriptional regulatory sequence is operably linked to a coding sequence if it controls the transcription of the coding sequence.

The term "or" is used herein to mean, and is used interchangeably with, the term "and/or", unless context clearly indicates otherwise.

The terms "proteins" and "polypeptides" are used interchangeably herein.

A "sample" includes any material that is obtained or prepared for detection of a molecular marker or a change in a molecular marker such as for example the methylation state, or any material that is contacted with a detection reagent or detection device for the purpose of detecting a molecular marker or a change in the molecular marker.

A "subject" is any organism of interest, generally a mammalian subject, such as a mouse, and preferably a human subject.

As used herein, the term "specifically hybridizes" refers to the ability of a nucleic acid probe/primer of the invention to hybridize to at least 12, 15, 20, 25, 30, 35, 40, 45, 50 or 100 consecutive nucleotides of a target sequence, or a sequence complementary thereto, or naturally occurring mutants thereof, such that it has less than 15%, preferably less than 10%, and more preferably less than 5% background hybridization to a cellular nucleic acid (e.g., mRNA or genomic DNA) other than the target gene. A variety of hybridization conditions may be used to detect specific hybridization, and the stringency is determined primarily by the wash stage of the hybridization assay. Generally high temperatures and low salt concentrations give high stringency, while low temperatures and high salt concentrations give low stringency. Low stringency hybridization is achieved by washing in, for example, about 2.0 x SSC at 50 °C, and high stringency is achieved with about 0.2 x SSC at 50 °C. Further descriptions of stringency are provided below.

As applied to polypeptides, "substantial sequence identity" means that two peptide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap which share at least 90 percent sequence identity, preferably at least 95 percent sequence identity, more preferably at least 99 percent sequence identity or more. Preferably, residue positions which are not identical differ by conservative amino acid substitutions. For example, the substitution of

amino acids having similar chemical properties such as charge or polarity are not likely to effect the properties of a protein. Examples include glutamine for asparagine or glutamic acid for aspartic acid.

As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., an HLTF polypeptide), which is partly or entirely heterologous (i.e., foreign) to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). An HLTF transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of a selected nucleic acid. An HLTF transgene can include a HLTF nucleotide sequence (SEQ ID NO: 3) or fragments thereof.

II. Overview

In certain aspects, the invention relates to methods for determining whether a patient is likely or unlikely to have a colon neoplasia. A colon neoplasia is any cancerous or precancerous growth located in, or derived from, the colon. The colon is a portion of the intestinal tract that is roughly three feet in length, stretching from the end of the small intestine to the rectum. Viewed in cross section, the colon consists of four distinguishable layers arranged in concentric rings surrounding an interior space, termed the lumen, through which digested materials pass. In order, moving outward from the lumen, the layers are termed the mucosa, the submucosa, the muscularis propria and the subserosa. The mucosa includes the epithelial layer (cells adjacent to the lumen), the basement membrane, the lamina propria and the muscularis mucosae. In general, the "wall" of the colon is intended to refer to the submucosa and the layers outside of the submucosa. The "lining" is the mucosa.

Precancerous colon neoplasias are referred to as adenomas or adenomatous polyps. Adenomas are typically small mushroom-like or wart-like growths on the

lining of the colon and do not invade into the wall of the colon. Adenomas may be visualized through a device such as a colonoscope or flexible sigmoidoscope. Several studies have shown that patients who undergo screening for and removal of adenomas have a decreased rate of mortality from colon cancer. For this and other reasons, it is generally accepted that adenomas are an obligate precursor for the vast majority of colon cancers.

When a colon neoplasia invades into the basement membrane of the colon, it is considered a colon cancer, as the term "colon cancer" is used herein. In describing colon cancers, this specification will generally follow the so-called "Dukes" colon cancer staging system. The characteristics that the describe a cancer are generally of greater significance than the particular term used to describe a recognizable stage. The most widely used staging systems generally use at least one of the following characteristics for staging: the extent of tumor penetration into the colon wall, with greater penetration generally correlating with a more dangerous tumor; the extent of invasion of the tumor through the colon wall and into other neighboring tissues, with greater invasion generally correlating with a more dangerous tumor; the extent of invasion of the tumor into the regional lymph nodes, with greater invasion generally correlating with a more dangerous tumor; and the extent of metastatic invasion into more distant tissues, such as the liver, with greater metastatic invasion generally correlating with a more dangerous disease state.

"Dukes A" and "Dukes B" colon cancers are neoplasias that have invaded into the wall of the colon but have not spread into other tissues. Dukes A colon cancers are cancers that have not invaded beyond the submucosa. Dukes B colon cancers are subdivided into two groups: Dukes B1 and Dukes B2. "Dukes B1" colon cancers are neoplasias that have invaded upto but not through the muscularis propria. Dukes B2 colon cancers are cancers that have breached completely through the muscularis propria. Over a five year period, patients with Dukes A cancer who receive surgical treatment (i.e. removal of the affected tissue) have a greater than 90% survival rate. Over the same period, patients with Dukes B1 and Dukes B2 cancer receiving surgical treatment have a survival rate of about 85% and 75%,

respectively. Dukes A, B1 and B2 cancers are also referred to as T1, T2 and T3-T4 cancers, respectively.

"Dukes C" colon cancers are cancers that have spread to the regional lymph nodes, such as the lymph nodes of the gut. Patients with Dukes C cancer who receive surgical treatment alone have a 35% survival rate over a five year period, but this survival rate is increased to 60% in patients that receive chemotherapy.

"Dukes D" colon cancers are cancers that have metastasized to other organs. The liver is the most common organ in which metastatic colon cancer is found. Patients with Dukes D colon cancer have a survival rate of less than 5% over a five year period, regardless of the treatment regimen.

In general, colon neoplasia develops through one of at least three different pathways, termed chromosomal instability, microsatellite instability, and the CpG island methylator phenotype (CIMP). Although there is some overlap, these pathways tend to present somewhat different biological behavior. By understanding the pathway of tumor development, the target genes involved, and the mechanisms underlying the genetic instability, it is possible to implement strategies to detect and treat the different types of colon neoplasias.

This application is based at least in part, on the recognition that certain target genes may be silenced or inactivated by the differential methylation of CpG islands in the 5'flanking or promoter regions of the target gene. CpG islands are clusters of cytosine-guanosine residues in a DNA sequence, that are prominently represented in the 5-flanking region or promoter region of about half the genes in our genome. In particular, this application is based at least in part on the recognition that differential methylation of the HLTF nucleotide sequence may be indicative of colon neoplasia. In one aspect, this application discloses that the HLTF gene can be a common target for methylation and epigenetic gene silencing in cancer cells (e.g., a colon neoplasia), and function as a candidate tumor suppressor gene.

HLTF (helicase-like transcription factor, also called HIP116a, Zbu1, RUSH1a and Smarca3) is a member of the SWI/SNF family. The SWI/SNF family

of genes encode members of multiprotein complexes that utilize the energy of ATP hydrolysis to alter nucleosome position or spacing (Muchardt, et al., 1999, J. Mol. Biol., 293:187-198; Sudarsanam, et al., 2000, Trends Genet., 16:345-351). HLTF has 5'- sequence-specific DNA-binding domains and can thus be targeted to specific promoters directly. For example, HLTF protein can bind to a promoter element (i.e., the B Box) of the plasminogen activator inhibitor-1 (PAI-1) gene and induce PAI-1 gene expression (Ding, et al., 1996, DNA Cell Biol. 15:429-442; Zhang, et al., 1997, Gene, 202:31-7). Functional interactions between Sp1 or Sp3 and HLTF were found to mediate basal expression from the PAI-1 gene (Ding, et al., 1999, J. Biol. Chem., 274:19573-19580). Recently, it has been found that HLTF is an activator of beta-globin transcription (Mahajan, et al., 2002, Blood, 99:348-56).

As noted above, early detection of colon neoplasia, coupled with appropriate intervention, is important for increasing patient survival rates. Present systems for screening for colon neoplasia are deficient for a variety of reasons, including a lack of specificity and/or sensitivity (e.g. Fecal Occult Blood Test, flexible sigmoidoscopy) or a high cost and intensive use of medical resources (e.g. colonoscopy). Alternative systems for detection of colon neoplasia would be useful in a wide range of other clinical circumstances as well. For example, patients who receive surgical and/or pharmaceutical therapy for colon cancer may experience a relapse. It would be advantageous to have an alternative system for determining whether such patients have a recurrent or relapsed colon neoplasia. As a further example, an alternative diagnostic system would facilitate monitoring an increase, decrease or persistence of colon neoplasia in a patient known to have a colon neoplasia. A patient undergoing chemotherapy may be monitored to assess the effectiveness of the therapy.

III. HLTF nucleic acids, polypeptides, and antibodies.

The present invention is based, at least in part, on the observation that HLTF nucleotide sequences are differentially methylated in certain HLTF-associated neoplasia, such as colon neoplasia. In one aspect, the application discloses HLTF nucleotide sequences having certain regions that are differentially methylated in

HLTF-associated neoplasia as set forth in SEQ ID NOs: 2-4, 21, and 39. In other embodiments, the application provides nucleotide sequences that are differentially methylated in HLTF-associated neoplasia as set forth in SEQ ID NOs: 4 and 21. Accordingly, in one embodiment the application provides isolated or recombinant nucleotide sequences that are at least 80%, 85%, 90%, 95%, 97%, 98%, 99% or 100% identical to a nucleic acid sequence as set forth in SEQ ID NOs: 2-4, 21, and 39. In yet other aspects, the application provides oligonucleotide sequences having at least 50, 75, or 100 consecutive base pairs of any one of the sequences as set forth in SEQ ID NOs: 2-4 and 21.

In certain alternative embodiments, the application provides the differentially methylated HLTF nucleotide sequence set forth in SEQ ID NOs: 4 and 21 and fragments thereof, wherein detection of methylation in any one of said fragments would be indicative of an HLTF-associated neoplasia such as colon neoplasia. One of ordinary skill in the art will appreciate that HLTF nucleic acid sequences complementary to SEQ ID NOs: 2-4 and 21, variants of SEQ ID NOs: 2-4 and 21 are also within the scope of this invention. Such variant nucleotide sequences include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as for example allelic variants.

In yet other embodiments, HLTF nucleotide sequences also include nucleotide sequences sequences that will hybridize under highly stringent conditions to nucleotide sequences designated in SEQ ID NOs: 2-4 and 21. As discussed above, one of ordinary skill in the art will understand readily that appropriate stringency conditions which promote DNA hybridization can be varied. One of ordinary skill in the art will understand readily that appropriate stringency conditions which promote DNA hybridization can be varied. For example, one could perform the hybridization at 6.0 x sodium chloride/sodium citrate (SSC) at about 45 °C, followed by a wash of 2.0 x SSC at 50 °C. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50 °C to a high stringency of about 0.2 x SSC at 50 °C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22 °C, to high stringency conditions at about 65 °C. Both temperature and salt

may be varied, or temperature or salt concentration may be held constant while the other variable is changed. In one embodiment, the invention provides nucleic acids which hybridize under low stringency conditions of 6 x SSC at room temperature followed by a wash at 2 x SSC at room temperature.

In yet another aspect, the application provides the methylated forms of nucleotide sequences as set forth in SEQ ID NOs: 4 and 21, wherein the cytosine bases of the CpG islands present in said sequences are methylated. In other words, the HLTF nucleotide sequences may be either in the methylated status (e.g., as seen in HLTF-associated neoplasias) or in the unmethylated status (e.g., as seen in normal cells). In further embodiments, the HLTF nucleotide sequences of the invention can be isolated, recombinant, and/or fused with a heterologous nucleotide sequence, or in a DNA library.

In addition to the differentially methylated HLTF nucleotide sequences, the present application discloses constitutively methylated HLTF nucleotide sequences (such as the Alu repeats and the "non-Alu constitutively methylated region" as set forth in Figure 10). Since such constitutively methylated HLTF nucleotide sequences are methylated in both normal cells and cancer cells, a person skilled in the art would appreciate the significance of detecting the differentially methylated HLTF nucleotide sequences as provided herein. Furthermore, although one of ordinary skill would expect the Alu's to be methylated, the finding of non-Alu constitutively methylated regions shows that one could not have computationally predicted the presence of the differentially methylated regions and their significance in the detection of HLTF-associated neoplasia.

In certain embodiments, the application contemplates any HLTF nucleotide sequence within the HLTF genomic sequence, SEQ ID NO: 39 (see Figure 28) that is differentially methylated in HLTF-associated neoplasia cells, but not in normal cells. Thus, assaying of the methylation status of such an HLTF nucleotide sequence can differentiate cancer cells from normal cells.

In certain embodiments, the present invention provides bisulfite-converted HLTF template DNA sequences as set forth in SEQ ID NOs: 5-8 and 22-25. Such

bisulfite-converted HLTF template DNA can be used for detecting the methylation status, for example, by an MSP reaction or by direct sequencing. In yet other embodiments, the bisulfite-converted HLTF nucleotide sequences of the invention also include nucleotide sequences that will hybridize under highly stringent conditions to any nucleotide sequence selected from SEQ ID NOs: 5-8 and 22-25.

In further aspects, the application provides methods for producing such bisulfite- converted nucleotide sequences, for example, the application provides methods for treating a nucleotide sequence with a bisulfite agent such that the unmethylated cytosine bases are converted to a different nucleotide base such as a uracil.

In yet other aspects, the application provides oligonucleotide primers for amplifying a region within the HLTF nucleic acid sequence of any one of SEQ ID NOs: 2-4 and 21. In certain aspects, a pair of the oligonucleotide primers (for example, SEQ ID NOs: 9-10) can be used in a detection assay, such as the HpaII assay. In certain aspects, primers used in an MSP reaction can specifically distinguish between methylated and non-methylated HLTF DNA, for example, SEQ ID NOs: 11-20, 26-38, and 40-43).

The primers of the invention have sufficient length and appropriate sequence so as to provide specific initiation of amplification of HLTF nucleic acids. Primers of the invention are designed to be "substantially" complementary to each strand of the HLTF nucleic acid sequence to be amplified. While exemplary primers are provided in SEQ ID NOs: 11-20, 26-38, and 40-43, it is understood that any primer that hybridizes with the bisulfite-converted HLTF sequence of any one of SEQ ID NOs: 2-4 and 21 are included within the scope of this invention and is useful in the method of the invention for detecting methylated nucleic acid, as described. Similarly, it is understood that any primers that would serve to amplify a methylation sensitive restriction site or sites within the differentially methylated region of SEQ ID NOs: 2-4 or 21 are included within the scope of this invention and is useful in the method of the invention for detecting nucleic methylated nucleic acid, as described.

The oligonucleotide primers of the invention may be prepared by using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidites are used as starting materials and may be synthesized as described by Beaucage, et al. (Tetrahedron Letters, 22:1859-1862, 1981). One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4,458,066.

The various Sequence Identification Numbers that have been used in this application are summarized below:

Table I

SEQ ID NO	Description/ Name	Corresponding Figure
1	amino acid sequence of human HLTF protein.	Figure 14.
2	5' genomic sequence of human HLTF gene, residues 1-3000.	Figure 15.
3	5' genomic sequence of human HLTF gene, residues 600-2600.	Figure 16.
4	5' genomic sequence of human HLTF gene, residues 1250-1800.	Figure 17.
5	methylated SEQ ID NO: 4, after bisulfite conversion.	Figure 18, top panel.
6	unmethylated SEQ ID NO: 4, after bisulfite conversion.	Figure 18, bottom panel.
7	complementary strand of SEQ ID NO: 5.	Figure 19, top panel.
8	complementary strand of SEQ ID NO: 6.	Figure 19, bottom panel.
9	P-HLTF1277F	Figure 20.
10	P-HLTF1724R	Figure 20.
11	P-HLTF1352MF	Figure 20.
12	P-HLTF1606MR	Figure 20.

13	P-HLTF1347UF	Figure 20.
14	P-HLTF1610UR	Figure 20.
15	P-HLTF1627MR	Figure 20.
16	P-HLTF1631UR	Figure 20.
17	P-HLTF1352MF(ASS)	Figure 20.
18	P-HLTF1607MR(ASS)	Figure 20.
19	P-HLTF1349UF(ASS)	Figure 20.
20	P-HLTF1611UR(ASS)	Figure 20.
21	5' genomic sequence of human HLTF gene, residues 1200-2500, sense strand.	Figure 21.
22	methylated SEQ ID NO: 21, after bisulfite conversion.	Figure 22.
23	unmethylated SEQ ID NO: 21, after bisulfite conversion.	Figure 23.
24	methylated antisense-strand of SEQ ID NO: 21, after bisulfite conversion.	Figure 24.
25	unmethylated antisense-strand of SEQ ID NO: 21, after bisulfite conversion.	Figure 25.
26	1P-HLTF1581MF(ASS)	Figure 27.
27	13P-HLTF1713MR(ASS)	Figure 27.
28	2P-HLTF1575UF(ASS)	Figure 27.
29	14P-HLTF1728UR(ASS)	Figure 27.
30	5P-HLTF1827MR(ASS)	Figure 27.
31	6P-HLTF1829UR(ASS)	Figure 27.
32	3P-HLTF1621MF	Figure 27.
33	7P-HLTF1873MR	Figure 27.
34	4P-HLTF1614UF	Figure 27.

35	8P-HLTF1878UR	Figure 27.
36	9P-HLTF1893MF(ASS)	Figure 27.
37	ALU(MB)2133FR(ASS)	Figure 27.
38	10P-HLTF1890UF(ASS)	Figure 27.
39	HLTF genomic sequence (GenBank accession No. NT_005616, complementary residues 572873-629300)	Figure 28.
40	15P-HLTF2201MF(ASS)	Figure 27
41	11P-HLTF2400MR(ASS)	Figure 27
42	16P-HLTF2197UF(ASS)	Figure 27
43	12P-HLTF2403UR(ASS)	Figure 27

In certain other aspects, the invention relates to HLTF nucleic acids that encode the HLTF polypeptide of SEQ ID NO: 1 and variants thereof. Variant include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and will, therefore, include coding sequences that differ from the nucleotide sequence of the coding sequence e.g., due to the degeneracy of the genetic code. In certain embodiments, variant nucleic acids will also include sequences that will hybridize under highly stringent conditions to a nucleotide sequence encoding SEQ ID NO: 1.

Isolated HLTF nucleic acids which differ from the nucleic acids encoding SEQ ID NO: 1 due to degeneracy in the genetic code are also within the scope of the invention. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC are synonyms for histidine) may result in "silent" mutations which do not affect the amino acid sequence of the protein. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject proteins will exist among mammalian cells. One skilled in the art will appreciate that these variations in one or more nucleotides (up to about 3-5% of the

nucleotides) of the nucleic acids encoding a particular protein may exist among individuals of a given species due to natural allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of this invention.

In certain embodiments, the recombinant HLTF nucleic acid may be operably linked to one or more regulatory nucleotide sequences in an expression construct. Regulatory nucleotide sequences will generally be appropriate for a host cell used for expression. Numerous types of appropriate expression vectors and suitable regulatory sequences are known in the art for a variety of host cells. Typically, said one or more regulatory nucleotide sequences may include, but are not limited to, promoter sequences, leader or signal sequences, ribosomal binding sites, transcriptional start and termination sequences, translational start and termination sequences, and enhancer or activator sequences. Constitutive or inducible promoters as known in the art are contemplated by the invention. The promoters may be either naturally occurring promoters, or hybrid promoters that combine elements of more than one promoter. An expression construct may be present in a cell on an episome, such as a plasmid, or the expression construct may be inserted in a chromosome. In a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selectable marker genes are well known in the art and will vary with the host cell used.

In certain aspects, the invention relates to HLTF polypeptide (SEQ ID NO: 1) described herein, and variants polypeptides thereof. In certain embodiments, variant polypeptides have an amino acid sequence that is at least 75% identical to an amino acid sequence as set forth in SEQ ID NO: 1. In other embodiments, the variant polypeptide has an amino acid sequence at least 80%, 85%, 90%, 95%, 97%, 98%, 99% or 100% identical to an amino acid sequence as set forth in SEQ ID NO: 1.

In certain aspects, variant HLTF polypeptides are agonists or antagonists of the HLTF polypeptide as set forth in SEQ ID NO: 1. Variants of these polypeptides

may have a hyperactive or constitutive activity, or, alternatively, act to prevent the tumor suppressor activity of HLTF. For example, a truncated form lacking one or more domain may have a dominant negative effect.

In certain aspects, isolated peptidyl portions of the HLTF polypeptide can be obtained by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding the polypeptide as set forth in SEQ ID NO: 1. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of the tumor suppressor function of HLTF.

In certain aspects, variant HLTF polypeptides containing one or more fusion domains. Well known examples of such fusion domains include, for example, polyhistidine, Glu-Glu, glutathione S transferase (GST), thioredoxin, protein A, protein G, and an immunoglobulin heavy chain constant region (Fc), maltose binding protein (MBP), which are particularly useful for isolation of the fusion polypeptide by affinity chromatography. For the purpose of affinity purification, relevant matrices for affinity chromatography, such as glutathione-, amylase-, and nickel- or cobalt- conjugated resins are used. Many of such matrices are available in "kit" form, such as the Pharmacia GST purification system and the QIAexpress™ system (Qiagen) useful with (HIS_6) fusion partners. Another fusion domain well known in the art is green fluorescent protein (GFP). This fusion partner serves as a fluorescent "tag" which allows the fusion polypeptide of the invention to be identified by fluorescence microscopy or by flow cytometry. The GFP tag is useful when assessing subcellular localization of the fusion HLTF polypeptide. The GFP tag is also useful for isolating cells which express the fusion HLTF polypeptide by flow cytometric methods such a fluorescence activated cell sorting (FACS). Fusion domains also include "epitope tags," which are usually short peptide sequences for which a specific antibody is available. Well known epitope tags for which specific monoclonal antibodies are readily available include FLAG, influenza virus haemagglutinin (HA), and c-myc tags. In some cases, the fusion domains have a

protease cleavage site, such as for Factor Xa or Thrombin, which allow the relevant protease to partially digest the fusion HLTF polypeptide and thereby liberate the recombinant polypeptide therefrom. The liberated polypeptide can then be isolated from the fusion partner by subsequent chromatographic separation.

Another aspect of the invention pertains to an isolated antibody specifically immunoreactive with an epitope of an HLTF polypeptide. For example, by using immunogens derived from an HLTF polypeptide (e.g., based on its cDNA sequences), anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, Antibodies: A Laboratory Manual ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal, such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the HLTF peptide. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of a polypeptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies.

In certain embodiment, antibodies of the invention may be useful as diagnostic or therapeutic agents for detecting or treating HLTF-associated diseases.

The term "antibody" as used herein is intended to include fragments thereof which are also specifically reactive with one of the HLTF polypeptide. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, $F(ab)_2$ fragments can be generated by treating antibody with pepsin. The resulting $F(ab)_2$ fragments can be treated to reduce disulfide bridges to produce Fab fragments. The antibody of the invention is further intended to include bispecific, single-chain, and chimeric and humanized molecules having affinity for the HLTF protein. In preferred embodiments, the antibody further comprises a label attached thereto and able to be detected, (e.g., the label can be a radioisotope, fluorescent compound, enzyme or enzyme co-factor).

IV. Assays and Drug Screening Methodologies

In certain aspects, the application provides assays and methods using the HLTF nucleotide sequences as molecular markers that distinguish between healthy cells and HLTF-associated diseased cells. For example, in one embodiment, the application provides methods and assays using the HLTF nucleotide sequences as markers that distinguish between healthy cells ands colon neoplasia cells. In one aspect, a molecular marker of the invention is a differentially methylated HLTF nucleotide sequence. In another aspect, another marker provided herein is the HLTF gene expression product.

In certain embodiments, the invention provides assays for detecting differentially methylated HLTF nucleotide sequences, such as the differential methylation patterns seen in any one of SEQ ID NOs: 2-4, 21 and 39, preferably, SED ID NOs: 4 and 21. Thus, a differentially methylated HLTF nucleotide sequence, in its methylated state, can be a HLTF-associated neoplasia-specific modification that serves as a target for detection using various methods described herein and the methods that are well within the purview of the skilled artisan in view of the teachings of this application.

In certain embodiments, methods of the present invention assaying for the methylation status of the HLTF nucleotide sequence in combination with one or more genes selected from HIC-1 (hypermethylated in cancer-1), p16, p14, TIMP-3, APC, PTEN, RARβ (retinoic acid receptor β), THBS1, hMLH1, and others. The present application provides that HLTF methylation correlates strongly with a pathway termed as the CpG island methylator phenotype (CIMP+), which may involve methylation of multiple genes, including p16, p14, HIC-1, TIMP-3, APC, PTEN, RARβ, THBS1, and hMLH1.

In certain aspects, such methods for detecting methylated HLTF nucleotide sequences are based on treatment of HLTF genomic DNA with a chemical compound which converts non-methylated C, but not methylated C (i.e., 5mC), to a different nucleotide base. One such compound is sodium bisulfite, which converts C, but not 5mC, to U. Methods for bisulfite treatment of DNA are known in the art

(Herman, et al., 1996, Proc Natl Acad Sci U S A, 93:9821-6; Herman and Baylin, 1998, Current Protocols in Human Genetics, N. E. A. Dracopoli, ed., John Wiley & Sons, 2:10.6.1-10.6.10; U.S. Patent No. 5,786,146). To illustrate, when an DNA molecule that contains unmethylated C nucleotides is treated with sodium bisulfite to become a compound-converted DNA, the sequence of that DNA is changed (C→U). Detection of the U in the converted nucleotide sequence is indicative of an unmethylated C.

The different nucleotide base (e.g., U) present in compound-converted nucleotide sequences can subsequently be detected in a variety of ways. In a preferred embodiment, the present invention provides a method of detecting U in compound-converted HLTF DNA sequences by using "methylation sensitive PCR" (MSP) (see, e.g., Herman, et al., 1996, Proc. Natl. Acad. Sci. USA, 93:9821-9826; U.S. Patent No. 6,265,171; U.S. Patent No. 6,017,704; U.S. Patent No. 6,200,756). In MSP, one set of primers (i.e., comprising a forward and a reverse primer) amplifies the compound-converted template sequence if C bases in CpG dinucleotides within the HLTF DNA are methylated. This set of primers is called "methylation-specific primers." Another set of primers amplifies the compound-converted template sequence if C bases in CpG dinucleotides within the HLTF 5' flanking sequence are not methylated. This set of primers is called "unmethylation-specific primers."

In methyl specific PCR the reactions use the compound-converted DNA from a sample in a subject. In assay for HLTF methylated DNA , methylation-specific primers are used. In the case where C within CpG dinucleotides of the target sequence of the DNA are methylated, the methylation-specific primers will amplify the compound-converted template sequence in the presence of a polymerase and an MSP product will be produced. If C within CpG dinucleotides of the target sequence of the DNA are not methylated, the methylation-specific primers will not amplify the compound-converted template sequence in the presence of a polymerase and an MSP product will not be produced

It is often also useful to run a control reaction for the detection of unmethylated HLTF DNA. The reactions uses the compound-converted DNA from a sample in a subject and unmethylation-specific primers are used. In the case where C within CpG dinucleotides of the target sequence of the DNA are unmethylated, the unmethylation specific primers will amplify the compound-converted template sequence in the presence of a polymerase and an MSP product will be produced. If C within CpG dinucleotides of the target sequence of the DNA are methylated, the unmethylation-specific primers will not amplify the compound-converted template sequence in the presence of a polymerase and an MSP product will not be produced. Note that a biologic sample will often contain a mixture of both neoplastic cells that give rise to a signal with methylation specific primers, and normal cellular elements that give rise to a signal with unmethylation-specific primers. The unmethyl specific signal is often of use as a control reaction, but does not in this instance imply the absence of colon neoplasia as indicated by the positive signal derived from reactions using the methylation specific primers.

Primers for an MSP reaction are derived from the compound-converted HLTF template sequence. Herein, "derived from" means that the sequences of the primers are chosen such that the primers amplify the compound-converted template sequence in an MSP reaction. Each primer comprises a single-stranded DNA fragment which is at least 8 nucleotides in length. Preferably, the primers are less than 50 nucleotides in length, more preferably from 15 to 35 nucleotides in length. Because the compound-converted HLTF template sequence can be either the Watson strand or the Crick strand of the double-stranded DNA that is treated with sodium bisulfite, the sequences of the primers is dependent upon whether the Watson or Crick compound-converted template sequence is chosen to be amplified in the MSP. Either the Watson or Crick strand can be chosen to be amplified.

The compound-converted HLTF template sequence, and therefore the product of the MSP reaction, can be between 20 to 3000 nucleotides in length, preferably between 50 to 500 nucleotides in length, more preferably between 80 to 150 nucleotides in length. Preferably, the methylation-specific primers result in an

MSP product of a different length than the MSP product produced by the unmethylation-specific primers.

A variety of methods can be used to determine if an MSP product has been produced in a reaction assay. One way to determine if an MSP product has been produced in the reaction is to analyze a portion of the reaction by agarose gel electrophoresis. For example, a horizontal agarose gel of from 0.6 to 2.0% agarose is made and a portion of the MSP reaction mixture is electrophoresed through the agarose gel. After electrophoresis, the agarose gel is stained with ethidium bromide. MSP products are visible when the gel is viewed during illumination with ultraviolet light. By comparison to standardized size markers, it is determined if the MSP product is of the correct expected size.

Other methods can be used to determine whether a product is made in an MSP reaction. One such method is called "real-time PCR." Real-time PCR utilizes a thermal cycler (i.e., an instrument that provides the temperature changes necessary for the PCR reaction to occur) that incorporates a fluorimeter (i.e. an instrument that measures fluorescence). The real-time PCR reaction mixture also contains a reagent whose incorporation into a product can be quantified and whose quantification is indicative of copy number of that sequence in the template. One such reagent is a fluorescent dye, called SYBR Green I (Molecular Probes, Inc.; Eugene, Oregon) that preferentially binds double-stranded DNA and whose fluorescence is greatly enhanced by binding of double-stranded DNA. When a PCR reaction is performed in the presence of SYBR Green I, resulting DNA products bind SYBR Green I and fluorescence. The fluorescence is detected and quantified by the fluorimeter. Such technique is particularly useful for quantification of the amount of the product in the PCR reaction. Additionally, the product from the PCR reaction may be quantitated in "real-time PCR" by the use of a variety of probes that hybridize to the product including TaqMan probes and molecular beacons. Quantitation may be on an absolute basis, or may be relative to a constitutively methylated DNA standard, or may be relative to an unmethyled DNA standard. In one instance the ratio of methylated HLTF derived product to unmethylated derived HLTF product may be constructed.

Methods for detecting methylation of the HLTF DNA in this invention are not limited to MSP, and may cover any assay for detecting DNA methylation. Another example method for detecting methylation of the HLTF DNA is by using "methylation-sensitive" restriction endonucleases. Such methods comprise treating the genomic DNA isolated from a subject with an methylation-sensitive restriction endonuclease and then using the restriction endonuclease-treated DNA as a template in a PCR reaction. Herein, methylation-sensitive restriction endonucleases recognize and cleave a specific sequence within the DNA if C bases within the recognition sequence are not methylated. If C bases within the recognition sequence of the restriction endonuclease are methylated, the DNA will not be cleaved. Examples of such methylation-sensitive restriction endonucleases include, but are not limited to HpaII, SmaI, SacII, EagI, MspI, BstUI, and BssHII. In this technique, a recognition sequence for a methylation-sensitive restriction endonuclease is located within the template DNA, at a position between the forward and reverse primers used for the PCR reaction. In the case that a C base within the methylation-sensitive restriction endonuclease recognition sequence is not methylated, the endonuclease will cleave the DNA template and a PCR product will not be formed when the DNA is used as a template in the PCR reaction. In the case that a C base within the methylation-sensitive restriction endonuclease recognition sequence is methylated, the endonuclease will not cleave the DNA template and a PCR product will be formed when the DNA is used as a template in the PCR reaction. Therefore, methylation of C bases can be determined by the absence or presence of a PCR product (Kane, et al., 1997, Cancer Res, 57:808-11). No sodium bisulfite is used in this technique.

Yet another exemplary method for detecting methylation of the HLTF DNA is called the modified MSP, which method utilizes primers that are designed and chosen such that products of the MSP reaction are susceptible to digestion by restriction endonucleases, depending upon whether the compound-converted template sequence contains CpG dinucleotides or UpG dinucleotides.

Yet other methods for detecting methylation of the HLTF DNA include the MS-SnuPE methods. This method uses compound-converted HLTF DNA as a

template in a primer extension reaction wherein the primers used produce a product, dependent upon whether the compound-converted template contains CpG dinucleotides or UpG dinucleotides (see e.g., Gonzalgo, et al., 1997, Nucleic Acids Res., 25:2529-31).

Another exemplary method for detecting methylation of the HLTF DNA is called COBRA (i.e., combined bisulfite restriction analysis). This method has been routinely used for DNA methylation detection and is well known in the art (see, e.g., Xiong, et al., 1997, Nucleic Acids Res, 25:2532-4).

In certain embodiments, the invention provides methods that involve directly sequencing the product resulting from an MSP reaction to determine if the compound-converted HLTF template sequence contains CpG dinucleotides or UpG dinucleotides. Molecular biology techniques such as directly sequencing a PCR product are well known in the art.

In alternative embodiments, the skilled artisan will appreciate that the present invention is based in part, on the recognition that HLTF functions as a tumor suppressor gene. Accordingly, in certain aspects, the invention provides assays for detecting molecular markers that distinguish between healthy cells and HLTF-associated diseases cells, such as colon neoplasia cells. As described above, one of the molecular markers of the present application includes that methylated HLTF nucleotide sequences. Thus, in one embodiment, assaying for the methylation status of the HLTF nucleotide sequence can be monitored for detecting an HLTF-silencing associated disease.

This application further provides another molecular marker : the HLTF gene expression transcript or the gene product. Thus, in another embodiment, expression of the HLTF nucleic acid or protein can be monitored for detecting an HLTF-silencing associated disease such as a colon neoplasia.

In certain embodiments, the invention provides detection methods by assaying the above-mentioned HLTF molecular markers so as to determine whether a patient has or does not have a disease condition. Further, such a disease condition

may be characterized by decreased expression of HLTF nucleic acid or protein described herein. In certain embodiments, the invention provides methods for determining whether a patient is or is not likely to have a HLTF-associated disease by detecting the expression of the HLTF nucleotide sequences. In further embodiments, the invention provides methods for determining whether the patient is having a relapse or determining whether a patient's cancer is responding to treatment.

In a preferred embodiment, the application provides method for detecting colon neoplasia. In certain embodiments, the present invention provides methods for detecting a colon neoplasia that is associated with silencing of HLTF gene. Such methods comprise assaying for the presence of a methylated HLTF nucleotide sequence in a sample obtained from a subject. In other aspects, the invention relates to methods for determining whether a patient is likely or unlikely to have a colon cancer. In further aspects, the invention relates to methods for monitoring colon neoplasia in a subject.

In certain embodiments, the invention provides assays for detecting HLTF protein or nucleic acid transcript described herein. In certain embodiments, a method of the invention comprises providing a biological sample and probing the biological sample for the HLTF expression which include protein or nucleic acid transcript of the HLTF. Information regarding the HLTF expression status, and optionally the quantitative level of the HLTF expression, may then be used to draw inferences about the nature of the biological sample and, if the biological sample was obtained from a subject, the health state of the subject.

In certain embodiments, methods of the present invention further comprise assaying for detecting a protein or a nucleic acid transcript selected from p16, THBS1, and hMLH1. The present inventors discovered that HLTF methylation correlates strongly with a pathway termed as the CpG island methylator phenotype (CIMP+), which may involve methylation of multiple genes, including p16, THBS1, and hMLH1. Methylation of these genes may lead to aberrant expression of the gene or the protein.

In certain embodiments, a method of the invention comprises detecting the presence of HLTF protein in a sample. Optionally, the method involves obtaining a quantitative measure of the HLTF protein in the sample. In view of this specification, one of skill in the art will recognize a wide range of techniques that may be employed to detect and optionally quantitate the presence of a protein. In preferred embodiments, HLTF protein is detected with an antibody. In many embodiments, an antibody-based detection assay involves bringing the sample and the antibody into contact so that the antibody has an opportunity to bind to proteins having the corresponding epitope. In many embodiments, an antibody-based detection assay also typically involves a system for detecting the presence of antibody-epitope complexes, thereby achieving a detection of the presence of the proteins having the corresponding epitope. Antibodies may be used in a variety of detection techniques, including enzyme-linked immunosorbent assays (ELISAs), immunoprecipitations, Western blots. Antibody-independent techniques for identifying a protein may also be employed. For example, mass spectroscopy, particularly coupled with liquid chromatography, permits detection and quantification of large numbers of proteins in a sample. Two-dimensional gel electrophoresis may also be used to identify proteins, and may be coupled with mass spectroscopy or other detection techniques, such as N-terminal protein sequencing. RNA aptamers with specific binding for the protein of interest may also be generated and used as a detection reagent.

Samples should generally be prepared in a manner that is consistent with the detection system to be employed. For example, a sample to be used in a protein detection system should generally be prepared in the absence of proteases. Likewise, a sample to be used in a nucleic acid detection system should generally be prepared in the absence of nucleases. In many instances, a sample for use in an antibody-based detection system will not be subjected to substantial preparatory steps. For example, urine may be used directly, as may saliva and blood, although blood will, in certain preferred embodiments, be separated into fractions such as plasma and serum.

In certain embodiments, a method of the invention comprises detecting the presence of an HLTF-expressed nucleic acid, such as an mRNA, in a sample. Optionally, the method involves obtaining a quantitative measure of the HLTF-expressed nucleic acid in the sample. In view of this specification, one of skill in the art will recognize a wide range of techniques that may be employed to detect and optionally quantitate the presence of a nucleic acid. Nucleic acid detection systems generally involve preparing a purified nucleic acid fraction of a sample, and subjecting the sample to a direct detection assay or an amplification process followed by a detection assay. Amplification may be achieved, for example, by polymerase chain reaction (PCR), reverse transcriptase (RT) and coupled RT-PCR. Detection of a nucleic acid is generally accomplished by probing the purified nucleic acid fraction with a probe that hybridizes to the nucleic acid of interest, and in many instances detection involves an amplification as well. Northern blots, dot blots, microarrays, quantitative PCR, and quantitative RT-PCR are all well known methods for detecting a nucleic acid in a sample.

In certain embodiments, the invention provides nucleic acid probes that bind specifically to an HLTF nucleic acid. Such probes may be labeled with, for example, a fluorescent moiety, a radionuclide, an enzyme or an affinity tag such as a biotin moiety. For example, the TaqMan® system employs nucleic acid probes that are labeled in such a way that the fluorescent signal is quenched when the probe is free in solution and bright when the probe is incorporated into a larger nucleic acid.

Immunoscintigraphy using monoclonal antibodies directed at the HLTF marker may be used to detect and/or diagnose a cancer. For example, monoclonal antibodies against the HLTF marker labeled with ^{99}Technetium, ^{111}Indium, ^{125}Iodine-may be effectively used for such imaging. As will be evident to the skilled artisan, the amount of radioisotope to be administered is dependent upon the radioisotope. Those having ordinary skill in the art can readily formulate the amount of the imaging agent to be administered based upon the specific activity and energy of a given radionuclide used as the active moiety. Typically 0.1-100 millicuries per dose of imaging agent, preferably 1-10 millicuries, most often 2-5 millicuries are administered. Thus, compositions according to the present invention

useful as imaging agents comprising a targeting moiety conjugated to a radioactive moiety comprise 0.1-100 millicuries, in some embodiments preferably 1-10 millicuries, in some embodiments preferably 2-5 millicuries, in some embodiments more preferably 1-5 millicuries.

In certain embodiments, the present invention provides drug screening assays for identifying test compounds which potentiate the tumor suppressor function of the HLTF gene. In one aspect, the assays detect test compounds which potentiate the expression level of the HLTF. In another aspect, the assays detect test compounds which inhibit the methylation of the HLTF nucleotide sequences. In certain embodiments, drug screening assays can be generated which detect test compounds on the basis of their ability to interfere with stability or function of the HLTF polypeptide. Alternatively, simple binding assays can be used to detect compounds that inhibit or potentiate the interaction between the HLTF polypeptide and its interacting protein (e.g., Sp1 or Sp3) or the binding of the HLTF polypeptide to a target DNA.

A variety of assay formats may be used and, in light of the present disclosure, those not expressly described herein will nevertheless considered to be within the purview of ordinary skill in the art. Assay formats can approximate such conditions as HLTF expression level, methylation status of HLTF sequence, tumor suppressing activity, transcriptional activating activity and may be generated in many different forms. In many embodiments, the invention provides assays including both cell-free systems and cell-based assays which utilize intact cells.

Compounds to be tested can be produced, for example, by bacteria, yeast or other organisms (e.g., natural products), produced chemically (e.g., small molecules, including peptidomimetics), or produced recombinantly. The efficacy of the compound can be assessed by generating dose-response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, the formation of complexes is quantitated in the absence of the test compound.

In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays of the present invention which are performed in cell-free systems, such as may be developed with purified or semi-purified proteins or with lysates, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the in vitro system, the assay instead being focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with other proteins or changes in enzymatic properties of the molecular target.

In certain embodiments, test compounds identified from these assays may be used in a therapeutic method for treating an HLTF-associated proliferative disease.

Still another aspect of the application provides transgenic non-human animals which express a heterologous HLTF gene, or which have had one or more genomic HLTF gene(s) disrupted in at least one of the tissue or cell-types of the animal. For instance, transgenic mice that are disrupted at their HLTF gene locus can be generated.

In another aspect, the application provides an animal model for an HLTF-associated proliferative disease, which has a mis-expressed HLTF allele. For example, a mouse can be bred which has an HLTF allele deleted, or in which all or part of one or more HLTF exons are deleted. Such a mouse model can then be used to study disorders arising from mis-expression of the HLTF gene.

Accordingly, the present application discloses transgenic animals which are comprised of cells (of that animal) containing an HLTF transgene and which preferably (though optionally) express an exogenous HLTF protein in one or more cells in the animal. The HLTF transgene can encode the wild-type form of the protein, or can encode homologs thereof, including both agonists and antagonists, as well as antisense constructs. The HLTF transgene can include an HLTF nucleotide

sequence (e.g., SEQ ID NOs: 2-4, 21 or 39) or fragments thereof. In preferred embodiments, the expression of the transgene is restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, cis-acting sequences that control expression in the desired pattern.

Genetic techniques which allow for the expression of transgenes can be regulated via site-specific genetic manipulation *in vivo* are known to those skilled in the art. For instance, genetic systems are available which allow for the regulated expression of a recombinase that catalyzes the genetic recombination a target sequence. As used herein, the phrase "target sequence" refers to a nucleotide sequence that is genetically recombined by a recombinase. The target sequence is flanked by recombinase recognition sequences and is generally either excised or inverted in cells expressing recombinase activity. Recombinase catalyzed recombination events can be designed such that recombination of the target sequence results in either the activation or repression of expression of the HLTF polypeptides. For example, excision of a target sequence which interferes with the expression of a recombinant HLTF gene can be designed to activate expression of that gene. This interference with expression of the protein can result from a variety of mechanisms, such as spatial separation of the HLTF gene from the promoter element or an internal stop codon. Moreover, the transgene can be made wherein the coding sequence of the gene is flanked recombinase recognition sequences and is initially transfected into cells in a 3' to 5' orientation with respect to the promoter element. In such an instance, inversion of the target sequence will reorient the subject gene by placing the 5' end of the coding sequence in an orientation with respect to the promoter element which allow for promoter driven transcriptional activation.

In an illustrative embodiment, either the *cre/loxP* recombinase system of bacteriophage P1 (Lakso et al., (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236; Orban et al., (1992) *Proc. Natl. Acad. Sci. USA* 89:6861-6865) or the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al., (1991) *Science* 251:1351-1355; PCT publication WO 92/15694) can be used to generate *in vivo* site-specific genetic recombination systems. Cre recombinase catalyzes the site-specific

recombination of an intervening target sequence located between *loxP* sequences. *loxP* sequences are 34 base pair nucleotide repeat sequences to which the Cre recombinase binds and are required for Cre recombinase mediated genetic recombination. The orientation of *loxP* sequences determines whether the intervening target sequence is excised or inverted when Cre recombinase is present (Abremski et al., (1984) *J. Biol. Chem.* 259:1509-1514); catalyzing the excision of the target sequence when the *loxP* sequences are oriented as direct repeats and catalyzes inversion of the target sequence when *loxP* sequences are oriented as inverted repeats.

V. Subjects and Samples

In certain aspects, the invention relates to a subject suspected of having or has an HLTF-associated disease such as colon neoplasia. Alternatively, a subject may be undergoing routine screening and may not necessarily be suspected of having such an HLTF-associated disease or condition. In a preferred embodiment, the subject is a human subject. and the HLTF associated disease is colon neoplasia.

Assaying for HLTF markers discussed above in a sample from subjects not known to have a colon neoplasia can aid in diagnosis of such a colon neoplasia in the subject. To illustrate, detecting the methylation status of the HLTF nucleotide sequence by MSP can be used by itself, or in combination with other various assays, to improve the sensitivity and/or specificity for detecting a colon neoplasia. Preferably, such a detection is made at an early stage in the development of cancer, so that treatment is more likely to be effective.

In addition to diagnosis, assaying of an HLTF marker in a sample from a subject not known to have colon neoplasia, can be prognostic for the subject (i.e., indicating the probable course of the disease). To illustrate, subjects having a predisposition to develop colon neoplasia may possess methylated HLTF nucleotide sequences. Assaying of HLTF markers in a samples from subjects can also be used to select a particular therapy or therapies which are particularly effective against the colon neoplasia in the subject, or to exclude therapies that are not likely to be effective.

Assaying of HLTF markers in samples from subjects that are known to have, or to have had, a cancer associated with silencing of the HLTF gene is also useful. For example, the present methods can be used to identify whether therapy is effective or not for certain subjects. One or more samples are taken from the same subject prior to and following therapy, and assayed for the HLTF markers. A finding that the HLTF marker is present in the sample taken prior to therapy and absent (or at a lower level) after therapy would indicate that the therapy is effective and need not be altered. In those cases where the HLTF marker is present in the sample taken before therapy and in the sample taken after therapy, it may be desirable to alter the therapy to increase the likelihood that the cancer will be eradicated in the subject. Thus, the present method may obviate the need to perform more invasive procedures which are used to determine a patient's response to therapy.

Cancers frequently recur following therapy in patients with advanced cancers. In this and other instances, the assays of the invention are useful for monitoring over time the status of an cancer associated with silencing of the HLTF gene. For subjects in which a cancer is progressing, an HLTF marker may be absent from some or all samples when the first sample is taken and then appear in one or more samples when the second sample is taken. For subjects in which cancer is regressing, an HLTF marker may be present in one or a number of samples when the first sample is taken and then be absent in some or all of these samples when the second sample is taken.

Samples for use with the methods described herein may be essentially any biological material of interest. For example, a sample may be a bodily fluid sample from a subject, a tissue sample from a subject, a solid or semi-solid sample from a subject, a primary cell culture or tissue culture of materials derived from a subject, cells from a cell line, or medium or other extracellular material from a cell or tissue culture, or a xenograft (meaning a sample of a cancer from a first subject, e.g., a human, that has been cultured in a second subject, e.g., an immuno-compromised mouse). The term "sample" as used herein is intended to encompass both a biological material obtained directly from a subject (which may be described as the

primary sample) as well as any manipulated forms or portions of a primary sample. A sample may also be obtained by contacting a biological material with an exogenous liquid, resulting in the production of a lavage liquid containing some portion of the contacted biological material. Furthermore, the term "sample" is intended to encompass the primary sample after it has been mixed with one or more additive, such as preservatives, chelators, anti-clotting factors, etc.

In certain embodiments, a bodily fluid sample is a blood sample. In this case, the term "sample" is intended to encompass not only the blood as obtained directly from the patient but also fractions of the blood, such as plasma, serum, cell fractions (e.g., platelets, erythrocytes, and lymphocytes), protein preparations, nucleic acid preparations, etc. In certain embodiments, a bodily fluid sample is a urine sample or a colonic effluent sample. In certain embodiments, a bodily fluid sample is a stool sample.

A subject is preferably a human subject, but it is expected that the molecular markers disclosed herein, and particularly their homologs from other animals, are of similar utility in other animals. In certain embodiments, it may be possible to detect an HLTF marker directly in an organism without obtaining a separate portion of biological material. In such instances, the term "sample" is intended to encompass that portion of biological material that is contacted with a reagent or device involved in the detection process.

In certain embodiments, DNA which is used as the template in an MSP reaction is obtained from a bodily fluid sample. Examples of preferred bodily fluids are blood, serum, plasma, a blood-derived fraction, stool, colonic effluent or urine. Other body fluids can also be used. Because they can be easily obtained from a subject and can be used to screen for multiple diseases, blood or blood-derived fractions are especially useful. For example, it has been shown that DNA alterations in colorectal cancer patients can be detected in the blood of subjects (Hibi, et al., 1998, Cancer Res, 58:1405-7). Blood-derived fractions can comprise blood, serum, plasma, or other fractions. For example, a cellular fraction can be prepared as a "buffy coat" (i.e., leukocyte-enriched blood portion) by centrifuging 5 ml of whole

blood for 10 min at 800 times gravity at room temperature. Red blood cells sediment most rapidly and are present as the bottom-most fraction in the centrifuge tube. The buffy coat is present as a thin creamy white colored layer on top of the red blood cells. The plasma portion of the blood forms a layer above the buffy coat. Fractions from blood can also be isolated in a variety of other ways. One method is by taking a fraction or fractions from a gradient used in centrifugation to enrich for a specific size or density of cells.

DNA is then isolated from samples from the bodily fluids. Procedures for isolation of DNA from such samples are well known to those skilled in the art. Commonly, such DNA isolation procedures comprise lysis of any cells present in the samples using detergents, for example. After cell lysis, proteins are commonly removed from the DNA using various proteases. RNA is removed using RNase. The DNA is then commonly extracted with phenol, precipitated in alcohol and dissolved in an aqueous solution.

VI. Therapeutic methods for HLTF-associated diseases.

Yet another aspect of this application pertains to methods of treating an HLTF-associated proliferative disease which arises from reduced expression or over-expression of the HLTF gene in cells. Such HLTF-associated proliferative diseases (for example, a colon neoplasia) can result from a wide variety of pathological cell proliferative conditions. In certain embodiments, treatment of an HLTF-associated proliferative disorder includes modulation of the HLTF gene expression or HLTF activity. The term "modulate" envisions the suppression of expression of HLTF when it is over-expressed, or augmentation of HLTF expression when it is under-expressed.

In an embodiment, the present invention provides a therapeutic method by using an HLTF gene construct as a part of a gene therapy protocol, such as to reconstitute the function of an HLTF protein (e.g., SEQ ID NO: 1) in a cell in which the HLTF protein is mis-expressed or non-expressed. To illustrate, cell types which exhibit pathological or abnormal growth presumably depend at least in part on a function of a HLTF protein. For example, gene therapy constructs encoding the

HLTF protein can be utilized in a colon neoplasia that is associated with silencing of the HLTF gene.

In certain embodiments, the invention provides therapeutic methods using agents which induce re-expression of HLTF. Loss of HLTF gene expression in an HLTF-associated diseased cells may be due at least in part to methylation of the HLTF nucleotide sequence, methylation suppressive agents such as 5-deoxyazacytidine or 5-azacytidine can be introduced into the diseased cells. Other similar agents will be known to those of skill in the art. A preferred embodiments the HLTF-associated disease is colon neoplasia associated with increased methylation of HLTF nucleotide sequences.

In certain embodiments, the invention provides therapeutic methods using a nucleic acid approach, for example, antisense nucleic acid, ribozymes or triplex agents, to block transcription or translation of a specific HLTF mRNA, either by masking that mRNA with an antisense nucleic acid or triplex agent or by cleaving it with a ribozyme. Such disorders include neurodegenerative diseases, for example. Antisense nucleic acids are DNA or RNA molecules that are complementary to at least a portion of a specific mRNA molecule (Weintraub, *Scientific American,* 262:40, 1990). In the cell, the antisense nucleic acids hybridize to the corresponding mRNA, forming a double-stranded molecule. The antisense nucleic acids interfere with the translation of the mRNA, since the cell will not translate an mRNA that is double-stranded. Antisense oligomers of about 15 nucleotides are preferred, since they are easily synthesized and are less likely to cause problems than larger molecules when introduced into a target HLTF over-producing cell. Use of an oligonucleotide to stall transcription is known as the triplex strategy since the oligomer winds around double-helical DNA, forming a three-strand helix. Therefore, these triplex compounds can be designed to recognize a unique site on a chosen gene (Maher, et al., *Antisense Res. and Dev.,* 1(3):227, 1991; Helene, C., *Anticancer Drug Design,* 6(6):569, 1991). Ribozymes are RNA molecules possessing the ability to specifically cleave other single-stranded RNA in a manner analogous to DNA restriction endonucleases. Through the modification of nucleotide sequences which encode these RNAs, it is possible to engineer molecules

that recognize specific nucleotide sequences in an RNA molecule and cleave it (Cech, *J. Amer. Med. Assn.*, 260:3030, 1988).

The present invention also provides gene therapy for the treatment of proliferative or immunologic disorders which are mediated by HLTF protein. Such therapy would achieve its therapeutic effect by introduction of the HLTF antisense polynucleotide into cells having the proliferative disorder. Alternatively, it may be desirable to introduce polynucleotides encoding full-length HLTF into diseased cells.

Delivery of antisense HLTF polynucleotide or the HLTF gene can be achieved using a recombinant expression vector such as a chimeric virus or a colloidal dispersion system. Especially preferred for therapeutic delivery of antisense sequences is the use of targeted liposomes. Various viral vectors which can be utilized for gene therapy as taught herein include adenovirus, herpes virus, vaccinia, or, preferably, an RNA virus such as a retrovirus. Preferably, the retroviral vector is a derivative of a murine or avian retrovirus. Examples of retroviral vectors in which a single foreign gene can be inserted include, but are not limited to: Moloney murine leukemia virus (MoMuLV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), and Rous Sarcoma Virus (RSV). Preferably, when the subject is a human, a vector such as the gibbon ape leukemia virus (GaLV) is utilized. A number of additional retroviral vectors can incorporate multiple genes. All of these vectors can transfer or incorporate a gene for a selectable marker so that transduced cells can be identified and generated. By inserting an HLTF sequence of interest into the viral vector, along with another gene which encodes the ligand for a receptor on a specific target cell, for example, the vector is target-specific. Retroviral vectors can be made target-specific by attaching, for example, a sugar, a glycolipid or a protein. Preferred targeting is accomplished by using an antibody to target the retroviral vector. Those skilled in the art will know of, or can readily ascertain without undue experimentation, specific polynucleotide sequences which can be inserted into the retroviral genome or attached to a viral envelope to allow target-specific delivery of the retroviral vector containing antisense HLTF polynucleotide or the HLTF gene.

The invention also relates to a medicament or pharmaceutical composition comprising an HLTF 5' flanking polynucleotide or an HLTF 5' flanking polynucleotide operably linked to the HLTF structural gene, respectively, in a pharmaceutically acceptable excipient or medium wherein the medicament is used for therapy of HLTF-associated cell proliferative disorders, such as a colon neoplasia.

Exemplification

The invention now being generally described, it will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present invention, and are not intended to limit the invention.

1. Amplification and Sequencing of HLTF cDNA.

HLTF cDNA sequence was obtained from Genbank (accession No. Z46606) and primers were chosen by using MacVector software (Oxford Molecular). The HLTF cDNA was amplified in two overlapping fragments (5' fragment 286F/1381R and 3' fragment 1317F/3456R). Primer sequences were as follows: HLTF-286F (5'-GCTCCTCTTGTCATCCCACTCA), HLTF-1381R (5'-CGTCTTTGCTTAGTCCATCTGCCTT), HLTF-1317F (5"-CGATGGTCTATGAAACTTGGA), and HLTF-3456R (5'-GAAATTGTGTCAGTAATACCTCTTCAC). The HLTF 5' genomic sequence was identified from Genbank genomic clone sequence (NT_005616).

2. Methylation-Specific PCR (MS-PCR).

500 ng DNA from each sample in a volume of 50 ul were denatured by NaOH (freshly made, final concentration, 0.2 M) at 37°C for 15 min. Next, 30 ul 10 mM hydroquinone (fresh) and 520 ul 3.0 M NaHSO4 (freshly prepared sodium bisulfite, pH5.0) were added, and incubated at 55°C for 16 hrs. Modified DNA was purified using Wizard DNA Clean-Up System (Promega). The reaction was desulphonated by NaOH at a final concentration of 0.3 M at room temperature for 15 min and neutralized by adding 10 M NH4OAc, pH7.0, to a final concentration of

3 M. DNA was precipitated with 3 volumes of absolute ethanol for 30 min at -80°C. The DNA pellet was then dissolved in distilled water to give approximately 10 ng/ul. Sodium bisulfite treated DNA was used as the template for subsequent methylation-specific PCR.

The primer sequences were based on the HLTF 5' genomic sequence and were specific for fully modified DNA. Primer set for the methylated DNA are P-HLTF1352MF: 5'-TGGGGTTTCGTGGTTTTTTCGCGC-3' and P-HLTF1606MR: 5'- CCGCGAATCCAATCAAACGTCGACG-3', which gives 254bp product. The primer set for the unmethylated DNA are P-HLTF1347UF: 5'-ATTTTTGGGGTTTTGTGGTTTTTTTGTGT -3' and P-HLTF1610UR: ATCACCACAAATCCAATCAAACATCAACA-3', which amplify 284bp fragment. PCR was carried out using a hot start at 95°C (9 minutes) and the following cycling parameters: 33 cycles of 95°C (45s), 66°C (45s), 72°C (45s), 72°C (5 minutes), and 4°C to cool. The PCR products were run on 3.0% agarose gel.

3. Cell culture and 5-Azacytidine treatment.

The cultures were grown and treated as described previously (Veigl, et al., 1998, Proc. Natl. Acad. Sci. USA, 95:8698-8702). The optimal tolerated doses were determined for each treated line, and two doses were used for some lines, ranging from 1 μg/ml to 3 μ g/ml.

4. Clonogenic assays and transfections.

Cells were plated in 6-well dishes (12,000-20,00 cells/well) 24 hours before transfection in a regular growing medium and transfected with 0.4 μg DNA/well with effectene (QIAGEN) according to manufacturers protocol. G418 (0.5mg/ml for FET, 0.6 mg/ml for V457)was added to the wells 48 hours after transfection, and cell were kept in G418 media (replaced bi-weekly) for 3 weeks, until tight colonies were observed. Colonies were stained with trypan blue and counted.

5. Statistical Methods.

Comparisons of HLTF methylation with sex, MLH1 methylation status, and CIMP (i.e., CpG island methylator phenotype) status were done using a two tailed Fisher's exact test. Comparison of HLTF methylation status with tumor site or stage was done using a Pearson's chi-squared test, with test for trend using a Mantel chi-squared test. Comparison of age distribution of smarca3 methylation in cancers and in normal tissue was done using a Wilcoxin non-parametric test.

6. HLTF is not mutated in colon cancers.

As several SWI/SNF family genes have been found to be altered in human cancers, we first determined the sequence of the HLTF cDNA amplified by RT-PCR from 34 colon cancer cell lines matched to primary patient samples in our colon cancer bank. Only one mutation was detected, a hemizygous nonsense mutation at codon 979. Thus HLTF is not a common target for gene mutation in colon cancer.

7. HLTF is frequently methylated and silenced in colon cancer cell lines.

In the process of HLTF sequence analysis in colon cancer cell lines we noted that 9 out of 34 of these cell lines did not express HLTF cDNA (Figure 1A). Southern analysis did not identify any alterations in the HLTF locus. Coincidentally, in some of the cell lines that had lost HLTF expression we previously had demonstrated silencing of the hMLH1 gene due to promoter methylation (Veigl, et al., 1998, Proc. Natl. Acad. Sci. USA, 95:8698-8702). We therefore examined the genomic sequence upstream of and within the HLTF gene (herein referred to as 5'- HLTF genomic sequence) which contained a CpG dense region that could potentially be methylated (Figure 2A). No TATA box consensus sequence was found within this region of 5' HLTF genomic sequence. However, it did contain a consensus initiator element, and two SP1 sites that are typical of TATA-less housekeeping gene promoters.

To test for methylation of this CpG-rich region, we used the technique of methylation specific PCR (Herman, et al., 1996, Proc. Natl. Acad. Sci. USA, 93:9821-9826), employing PCR primers specific for amplification of either methylated or unmethylated DNA templates (Figure 2A). As shown in Figure 2B,

all colon cancer cell lines that lacked HLTF gene expression demonstrated methylation of CpG sites within the 5' HLTF genomic sequence; whereas, methylation was not detected in the HLTF expressing cell lines. These results were confirmed by two independent MS-PCR assays that tested different HLTF CpG sites, as well as by resistance of the 5' HLTF genomic sequence to digestion with a methylation-sensitive restriction enzyme, HpaII enzyme. Thus, cell lines that had silenced the HLTF gene demonstrated methylation across the CpG sites within this entire region, whereas HLTF-expressing cell lines assayed as free of methylated CpG sites.

For three of these HLTF methylated cell lines, DNA from matched normal and antecedant tumor DNA was additionally available (V6, V8, and V432). In each of these cases, HLTF DNA methylation was detected in the primary tumors, but was absent in the matched normal tissues (Figure 2C), verifying that HLTF methylation and silencing was a true somatic event and was not an artifact of cell line cultures.

8. Re-induction of HLTF expression.

To establish that methylation was responsible for silencing HLTF gene expression, cell lines with HLTF DNA methylation were treated with 5-azacytidine (5-azaC), a demethylating agent. As shown in Figure 1B, 5-azaC treatment reactivated HLTF expression in all these cell lines, though Vaco457 required higher dose of 5-azaC for reactivation, compared to other cell lines. However, 5-AzaC did not further increase HLTF expression in control cell lines in which HLTF expression was constitutive and in which the basal HLTF DNA was unmethylated.

9. HLTF methylation is widespread in primary colon cancer.

To further establish the frequency of HLTF methylation in primary colon cancer tumors, we analyzed 63 pairs of primary colon tumors along with matched normal tissues (Figure 3). HLTF methylation was detected in 27 of 63 (45%) colon cancer cases. In contrast, no evidence of HLTF gene silencing was detected in 30 lung tumor cell lines, or 8 breast cancer samples tested.

The finding of HLTF methylation in colon cancer tumors and cell lines was not correlated with patients' sex (p=0.31) or with age (p=0.14) (Figure 4A), with a median age of 72 in persons with HLTF methylated cancers versus 68 in those with HLTF unmethylated cancers. In the overwhelming majority of cases (84%), HLTF methylation was detected only in the colon cancers, and was absent from the same individuals' normal colon tissues. HLTF methylation thus substantially arose in these individuals specifically during the neoplastic process. However, in 16% of individuals whose colon cancers demonstrated HLTF methylation, very faint HLTF methylation was also detectable in histologically normal colon tissue.

This was also the case in a small subset of all individuals from whom normal colon tissues were available to us, among whom faint HLTF methylation was detectable in 9 out of 78 normal colon samples (12%). Normal colon tissues showing faint HLTF methylation were in general those derived from the oldest individuals studied, with a median age of 81 for individuals demonstrating faint HLTF methylation in normal colon tissues versus 67 for those with whose normal colon tissue showed only unmethylated HLTF (p=0.02) (Figure 4B). We cannot determine whether HLTF methylated cells when detected in normal colon tissues are derived from contamination by cancer cells derived from a frank cancer that was concurrently resected in the same colon specimen from which our normal sample was taken, are derived from microscopic early colon neoplasms, or alternatively whether in some instances HLTF methylation can be initiated in the aging colon separately from neoplasia, either as a stochastic event, or in response to endogenous or exogenous genotoxins. However, we favor the explanation that these signals arise from contaminating cancer cells, as no HLTF methylation was detected in colons from any of 12 individuals undergoing colon resection for non-malignant disease. Compared to HLTF expressing cancers, cancers with methylated and silenced HLTF alleles showed a borderline significant trend (p=0.06) to be more likely to arise in the proximal right colon and less likely to arise in the left colon or rectum, a trend similar to that previously observed for HNPCC and sporadic MSI colon cancers (Kinzler, et al., 1996, Cell, 87:159-170) (Figure 5A vs. 5B). The distribution by tumor stage (adenoma; Dukes' stage B, C or D cancer primary; or metastatic lesion) was also significantly different between HLTF methylated and

non- methylated colon neoplasms (p=0.02). An *a postiori* grouping of the tumors into a non-metastatic subset and a metastatic subset (Dukes D primary cancers or cancers from distant metastatic sites) suggests the hypothesis that this is due to a lesser likelihood of HLTF methylated tumors being metastastic (nominal p value =0.01) (Figure 5C vs. 5D).

To determine the timing of onset of HLTF silencing during colon carcinogenesis, we additionally analyzed a group of 14 early and late adenomas for HLTF CpG island methylation. HLTF methylation was detected in 3 of the adenomas tested, all of which were greater than 1.5 cm in size, suggesting that HLTF methylation can occur as early as the late adenoma stage of colon neoplasia. Detection of HLTF methylation may thus be of value for detecting the early and most curable stages of colon neoplasia.

10. HLTF methylation defines a singular group of colon cancers.

Recently, it has been suggested that certain colon cancers are typified by a high frequency of gene promoter methylation and represent a distinct pathway termed the CpG island methylator phenotype (CIMP+) (Toyota, et al., 1999, Proc. Natl. Acad. Sci. USA, 96:8681-8686; Toyota, et al., 2000, Proc. Natl. Acad. Sci. USA, 97:710-715). Tumors exhibiting this phenotype (CIMP+) show concordant CpG island methylation affecting multiple genes, including hMLH1, p16, and THBS1. To establish whether HLTF methylation correlates with hMLH1 methylation, and/or with the CIMP+ phenotype, 87 colon cancer cases examined for HLTF methylation were also examined for hMLH1 methylation, and 64 were further assayed for CIMP+ or CIMP- phenotype as determined by methylation status of MINT1, MINT2, MINT31, and MINT27 locii (Toyota, et al., 1999, Proc. Natl. Acad. Sci. USA, 96:8681-8686). HLTF methylation correlated with CIMP+ phenotype ($p<0.001$) (Figure 6A) and as well with hMLH1 gene methylation ($p<0.0001$) (Figure 6B). However, HLTF-methylated tumors essentially defined a distinct subclass of colon cancers that did not fall exclusively into either the hMLH1 methylated or CIMP+ groups.

To further determine whether HLTF methylation defines a singular group of colon cancers, we used restriction landmark genomic scanning (RLGS) analysis (Costello, et al., 2000, Nat. Genet., 24:132-138) to compare the patterns of global genome methylation in a group of twelve colon cancer cell lines, six of which demonstrated HLTF methylation and silencing and in six of which HLTF was unmethylated and expressed. 497 loci demonstrated methylation present in at least one of the 12 colon cancer cell lines.

However, none of these loci demonstrated the presence of methylation across the six colon cancers in which HLTF was methylated and silenced, as well as the absence of methylation across the 6 colon cancers expressing an unmethylated HLTF allele. While RLGS analysis samples only a portion of the genome, this data independently suggests that HLTF methylation is a unique event, and does not necessarily reflect a genome-wide increase in promoter methylation.

11. HLTF reconstitution induces growth suppression.

The high frequency of HLTF methylation observed in colon cancer suggested that inactivation of this gene might confer a selective advantage. To assay for such an advantage we examined the effect of HLTF transfection on colony formation in three HLTF methylated and non-expressing colon cancer cell lines (V457, V8-2, RCA) as compared to three HLTF unmethylated and expressing colon cancer cell lines (FET, V364, V429). Reconstitution of HLTF expression in HLTF methylated cells suppressed colony forming ability by 75% in each of the three lines tested ($p<0.0001$ for each) (Figure 7A). In contrast, transfection of HLTF did not show significant colony suppression in the any of the three cell lines that already expressed endogenous HLTF (Figure 7B). Growth suppression by exogenous HLTF was thus specific to colon cancers that had silenced the endogenous alleles ($p<0.01$ for the difference in effect of HLTF transfection in HLTF methylated versus unmethylated cell lines). Transient transfections showed both the HLTF methylated and unmethylated cells were well able to express exogenous HLTF, as determine by Western analysis for a V5 epitope tag attached to HLTF in the expression vector

(Figure 7C). These findings suggest that HLTF methylation and silencing indeed confers a growth advantage in a distinct subclass of colon cancers.

In sum, certainly future studies can be expected to further elucidate the presumptive pathogenetic role that we suggest for HLTF inactivation in colon cancer. Moreover, the high frequency of HLTF methylation in colon cancer may also be useful in potential translational applications. We and others have shown that methylated promoter DNA can be detected in the blood of some cancer patients (Grady, et al., 2001, Cancer Res., 61:900-902). Thus, it will also be attractive to explore the possibility that assays for methylation of HLTF in body fluids may be of future value for early detection of colon cancer incidence, relapse or prognosis.

Incorporation by Reference

All publications and patents mentioned herein are hereby incorporated by reference in their entirety as if each individual publication or patent was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control.

Equivalents

While specific embodiments of the subject invention have been discussed, the above specification is illustrative and not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of this specification and the claims below. The full scope of the invention should be determined by reference to the claims, along with their full scope of equivalents, and the specification, along with such variations.

We claim:

1. A method for detecting differential methylation patterns in an HLTF nucleotide sequence, comprising:

a) obtaining a sample from a patient;

b) assaying said sample for the presence of methylation within a nucleotide sequence as set forth in any one of SEQ ID NOs: 2-4, 21 or 39;

c) obtaining a sample from a healthy subject;

d) assaying for the presence of methylation in a nucleotide sequence as set forth in any one of SEQ ID NOs: 2-4, 21 or 39; and

e) comparing the methylation patterns in the sample from the subject to the methylation patterns in the normal sample.

2. A method for detecting an HLTF-associated neoplasia, comprising:

a) obtaining a sample from a patient; and

b) assaying said sample for the presence of methylation within a nucleotide sequence as set forth in any one of SEQ ID NOs: 4 or 21;

wherein methylation of said nucleotide sequence is indicative of an HLTF-associated neoplasia.

3. The method of any one of claims 1 or 2, wherein the sample is a bodily fluid selected from the group consisting of blood, serum, plasma, a blood-derived fraction, stool, urine, and a colonic effluent.

4. The method of claim 3, wherein the bodily fluid is obtained from a subject suspected of having or is known to have an HLTF-associated neoplasia.

5. The method of claim 4, wherein said HLTF-associated neoplasia is colon neoplasia.

6. The method of any one of claims 1 or 2, comprising assaying for the presence of methylation within the HLTF sequence of SEQ ID NO: 4.

7. The method of any of claims 1-6, wherein the assay is methylation-specific PCR.

8. The method of claim 7, comprising:

a) treating DNA from the sample with a compound that converts non-methylated cytosine bases in the DNA to a different base;

b) amplifying a region of the compound converted HLTF nucleotide sequence with a forward primer and a reverse primer; and

c) analyzing the methylation patterns of said HLTF nucleotide sequences.

9. The method of claim 7, comprising:

a) treating DNA from the sample with a compound that converts non-methylated cytosine bases in the DNA to a different base;

b) amplifying a region of the compound converted HLTF nucleotide sequence with a forward primer and a reverse primer; and

c) detecting the presence and/or amount of the amplified product.

10. The method of claim 7, wherein the forward primers are selected from SEQ ID NOs: 13, 17, 19, 26, 28, 32, 34, 36, 38, 40, and 42.

11. The method of claim 7, wherein the reverse primers are selected from SEQ ID NOs: 14, 15, 16, 18, 20, 27, 29, 30, 31, 33, 35, 37, 41, and 43.

12. The method of claim 7, wherein the compound used to treat DNA is a bisulfite compound.

13. The method of any of claims 1-6, wherein the assay comprises using a methylation-specific restriction enzyme.

14. The method of claim 11, wherein said methylation-specific restriction enzyme is selected from HpaII, SmaI, SacII, EagI, MspI, BstUI, and BssHII.

15. The method of claim 14, wherein the primers are SEQ ID Nos 11 and 12.

16. A method for detecting an HLTF-associated neoplasia in a subject, comprising detecting HLTF protein or nucleic acid expression in a sample from the subject.

17. The method of claim 16, wherein the sample is a bodily fluid selected from the group consisting of blood, serum, plasma, a blood-derived fraction, stool, urine, and a colonic effluent.

18. The method of claim 17, wherein the bodily fluid is from a subject suspected of having or known to have an HLTF-associated neoplasia.

19. The method of claim 18, wherein the HLTF-associated neoplasia is colon neoplasia.

20. The method of claim 16, wherein the HLTF protein is detected by immunoassays.

21. A method for identifying an agent which enhances HLTF protein or nucleic acid expression in a diseased cell associated with HLTF gene silencing, comprising:

a) contacting the cell with a sufficient amount of the agent under suitable conditions;

b) quantitatively determining the amount of HLTF protein or nucleic acid; and

c) comparing the amount of HLTF protein or nucleic acid with the amount of HLTF protein or nucleic acid in the absence of the agent,

wherein a greater amount of HLTF protein or nucleic acid in the presence of the agent than in the absence of the agent indicates that the agent enhances HLTF protein or nucleic acid expression.

22. The method of claim 21, wherein said HLTF gene silencing is due to differential methylation of an HLTF nucleotide sequence.

23. The method of claim 22, wherein differential methylation occurs within an HLTF nucleotide sequence set forth in any one of SEQ ID NOs: 2-4, 21 and 39.

24. The method of claim 21, wherein the diseased cell is from a subject having colon neoplasia.

25. A method for monitoring over time a HTLF-associated neoplasia comprising:

a) detecting the methylation status of an HLTF nucleotide sequence in a sample from the subject for a first time; and

b) detecting the methylation status of the HLTF nucleotide sequence in a sample from the same subject at a later time;

wherein absence of methylation in the HLTF nucleotide sequence taken at a later time and the presence of methylation in the HLTF nucleotide sequence taken at the first time is indicative of cancer regression.

26. The method of claim 25, wherein the sample is a bodily fluid selected from the group consisting of blood, serum, plasma, a blood-derived fraction, stool, urine, and a colonic effluent.

27. The method of claim 25, wherein the HTLF-associated neoplasia is colon neoplasia.

28. A method for treating an HLTF-associated proliferative disease in a subject, comprising administering to the subject a sufficient amount of a compound, wherein the compound modulates the HLTF protein or nucleic acid expression.

29. The method of claim 28, wherein the disease is associated with methylation of an HLTF nucleic acid sequence, and the compound induces HLTF expression.

30. The method of claim 29, the compound is a demethylation agent selected from 5-azacytidine and 5-deoxy-azacytidine.

31. The method of claim 28, wherein the HLTF-associated proliferative disease is a colon neoplasia.

32. A method for treating an HLTF-associated neoplasia in a subject, comprising administering to the subject a vector containing an HLTF nucleic acid which is operably linked to a heterologous promoter.

33. The method of claim 32, wherein the HLTF nucleic acid encodes a polypeptide at least 90% identical to SEQ ID NO: 1.

34. The method of claim 32, wherein the cancer is a colon neoplasia.

35. A bisulfite-converted methylated HLTF nucleotide sequence selected from the group consisting of:

a nucleotide sequence of any one of SEQ ID NOs: 5-8 or a fragment thereof;

a complement of any one of SEQ ID Nos: 5-8; and

a nucleotide sequence that hybridizes under stringent conditions to a nucleotide sequence of any one of SEQ ID NOs: 5-8.

36. A bisulfite-converted methylated HLTF nucleotide sequence selected from the group consisting of:

a nucleotide sequence of any one of SEQ ID NOs: 22-25 or a fragment thereof;

a complement of any one of SEQ ID Nos: 22-25; and

a nucleotide sequence that hybridizes under stringent conditions to a nucleotide sequence of any one of SEQ ID NOs: 22-25.

37. Oligonucleotide primers for detecting methylation of an HLTF nucleotide sequence, selected from SEQ ID NOs: 9-20, and 26-29.

38. A kit for detecting an HLTF-associated neoplasia in a subject, comprising at least two primers of claim 37.

39. The kit of claim 38, further comprising a compound to convert a template DNA.

40. The kit of claim 41, wherein the compound is bisulfite.

41. The kit of claim 40, wherein each primer comprises at least a CpG dinucleotide.

42. A method of converting a nucleic acid sequence at least 95% identical to and one of SEQ ID NOs: 4 and 21 to a bisulfite converted sequence comprising:

a) providing a nucleotide acid having a nucleotide sequence as set forth in any one of SEQ ID NOs: 4 and 21; and

b) adding a bisulfite compound;

whereby the unmethylated cytosine bases of the CpG islands are converted to a different base.

43. The method of claim 42, wherein the unmethylated cytosine is converted to a uracil.

44. A nucleic acid sequence as prepared by the method of claim 44.

45. An isolated or recombinant methylated HLTF nucleic acid, comprising a nucleotide sequence of any one of SEQ ID NOs: 2-4 and 21, wherein the cytosine of the CpG island is methylated.

46. An isolated or recombinant HLTF nucleic acid, selected from the group consisting of:

a nucleotide sequence of any one of SEQ ID NOs: 2-4 and 21 or a fragment thereof;

a complement of any one of SEQ ID Nos: 2-4 and 21;

a nucleotide sequence that hybridizes under stringent conditions to a nucleotide sequence of any one of SEQ ID NOs: 2-4 and 21;

a nucleotide sequence that is at least 98% identical to the nucleotide sequence of any one of SEQ ID NOs: 2-4 and 21;

an nucleotide sequence comprising at least 50 consecutive base pairs of any one SEQ ID Nos; 2-4 and 21;

wherein the HLTF nucleotide sequence is differentially methylated in an HLTF-associated disease cell.

Figure 1

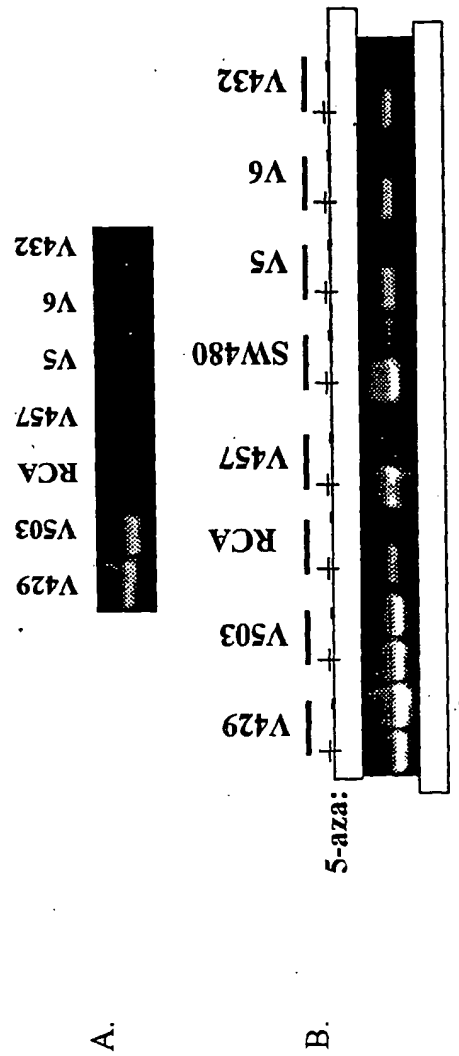

V432
V6
V5
SW480
V457
RCA
V503
V429
5-aza:
V432
V6
V5
V457
RCA
V503
V429
A.
B.

Figure 2

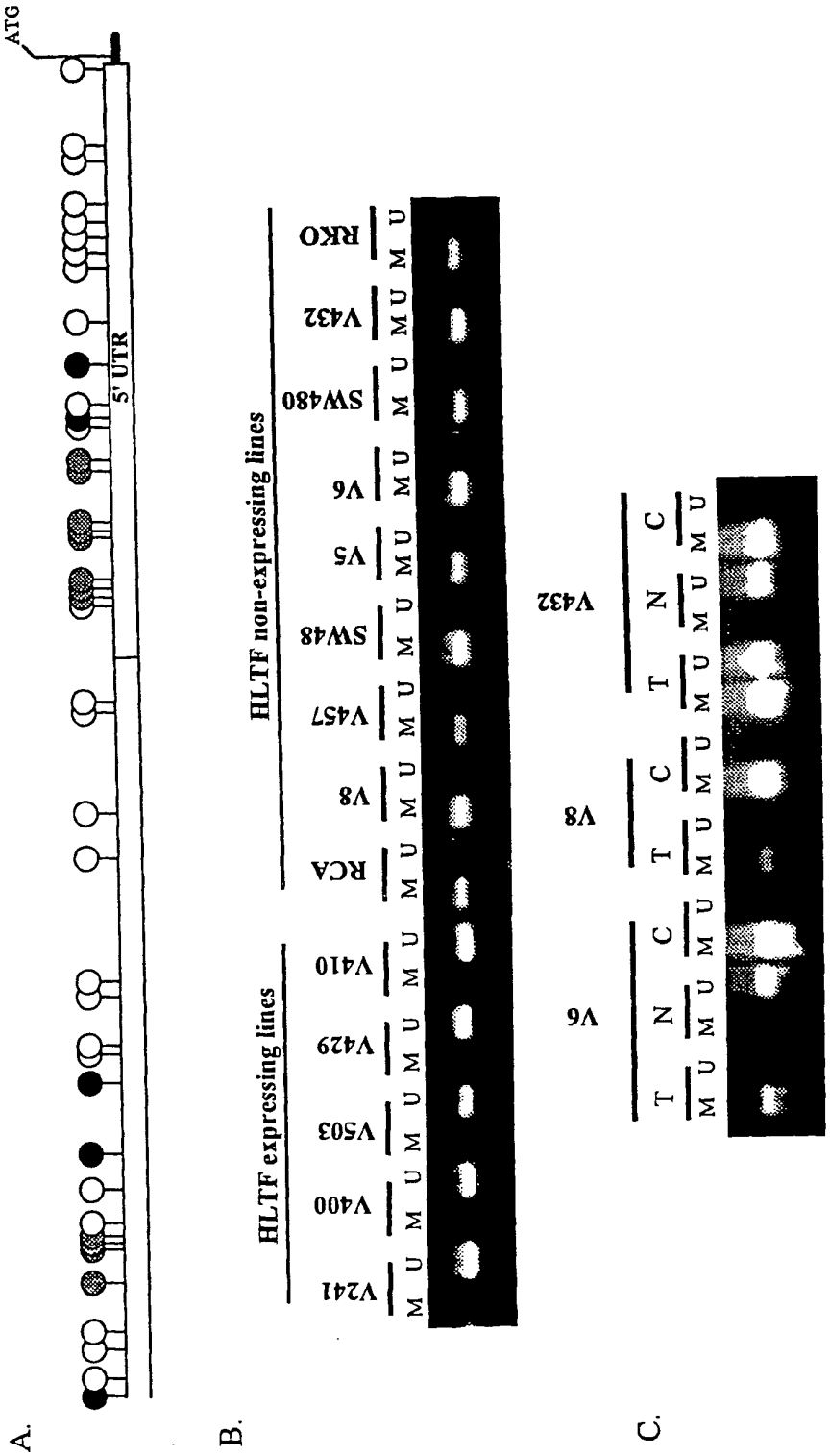

A.
ATG
5' UTR
B.
HLTF expressing lines
HLTF non-expressing lines
V241
V400
V503
V429
V410
RCA
V8
V457
SW48
V5
V6
SW480
V432
RKO
M U M U M U M U M U M U M U M U M U M U
C.
V6
V8
V432
T N C T N C
M U M U M U M U M U M U

Figure 3

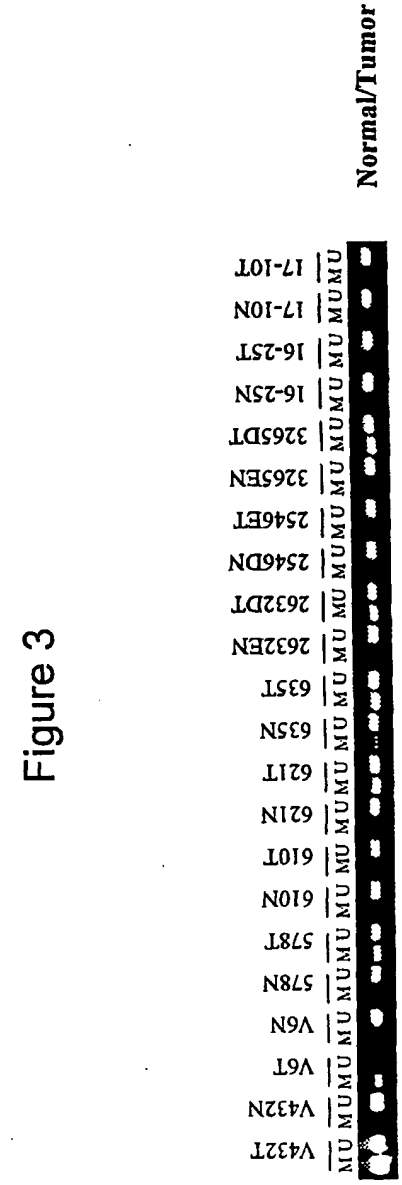

Normal/Tumor
V432T
V432N
V6T
V6N
578N
578T
610N
610T
621N
621T
635N
635T
2632EN
2632DT
2546DN
2546ET
3265EN
3265DT
16-25N
16-25T
17-10N
17-10T

Figure 4

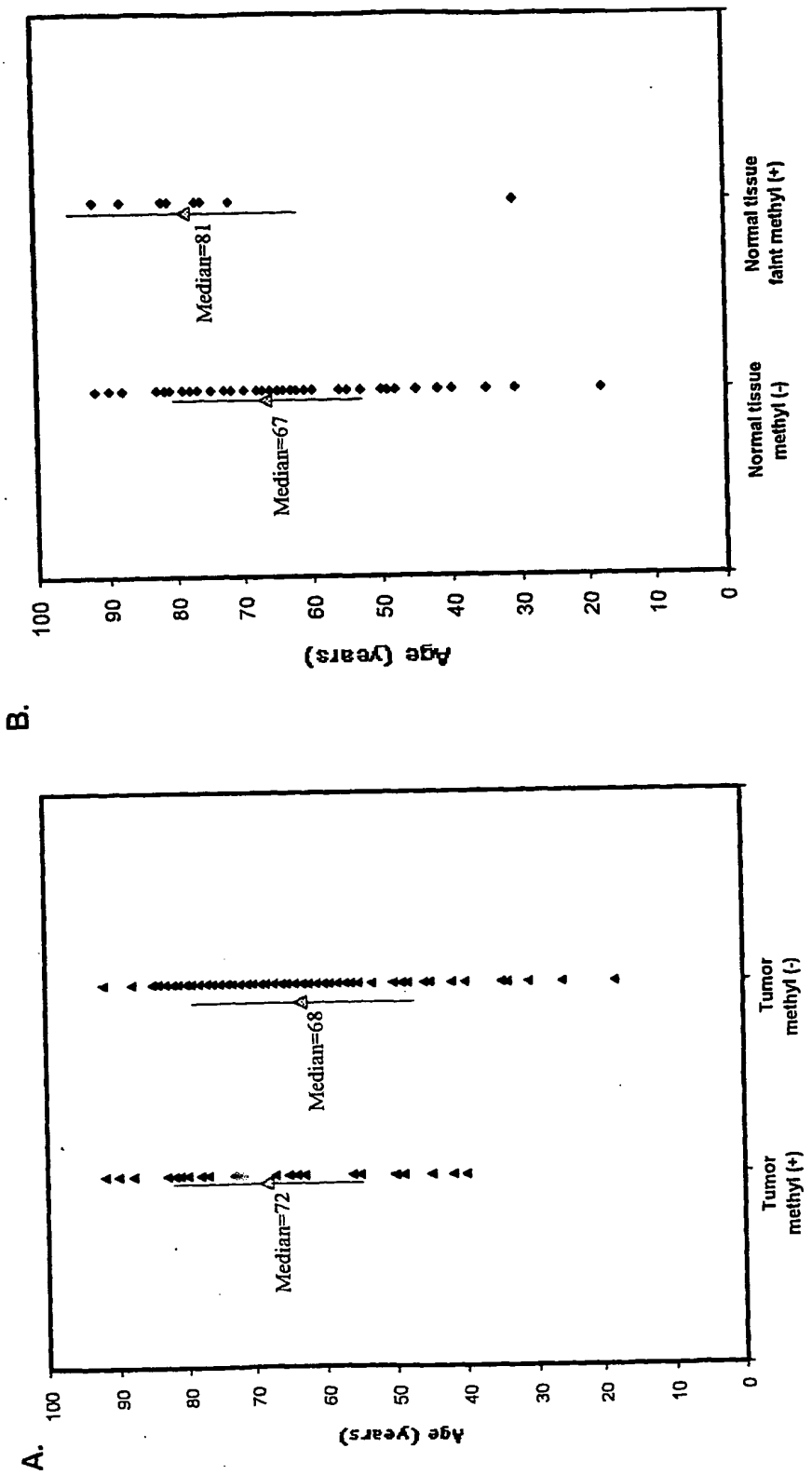

B.
Age (years)
100
90
80
70
60
50
40
30
20
10
0
Median=81
Median=67
Normal tissue faint methyl (+)
Normal tissue methyl (-)
A.
Age (years)
100
90
80
70
60
50
40
30
20
10
0
Median=68
Median=72
Tumor methyl (-)
Tumor methyl (+)

Figure 5

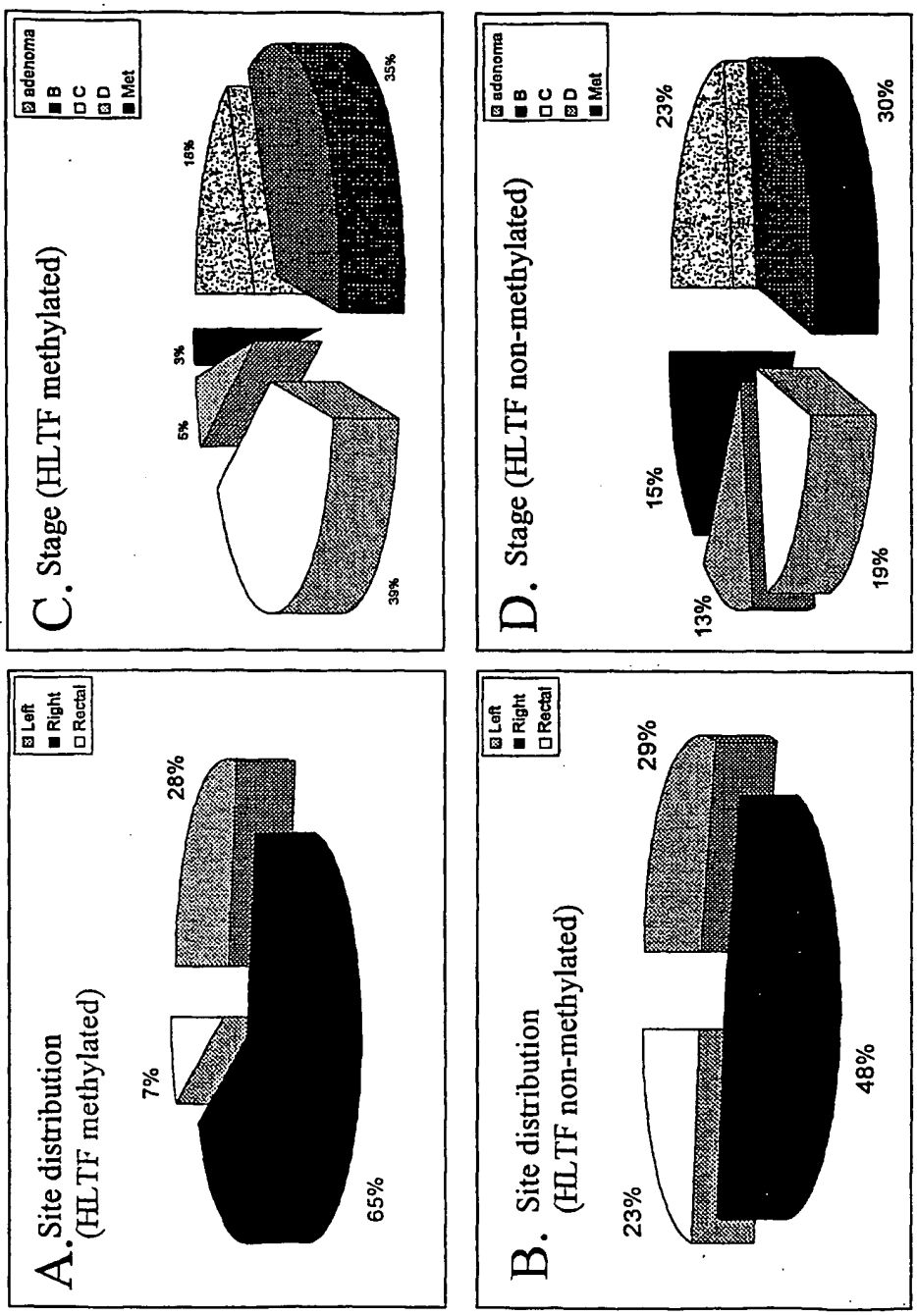

A. Site distribution (HLTF methylated)
Left
Right
Rectal
7%
28%
65%
B. Site distribution (HLTF non-methylated)
Left
Right
Rectal
23%
29%
48%
C. Stage (HLTF methylated)
adenoma
B
C
D
Met
6%
3%
18%
35%
39%
D. Stage (HLTF non-methylated)
adenoma
B
C
D
Met
15%
23%
30%
13%
19%

Figure 6

A.

CIMP status	HLTF methylated	HLTF nonmethylated
CIMP+	20	11
CIMP-	8	25

B.

	HLTF methylated	HLTF nonmethylated
MLH1 methylated	20	6
MLH1 nonmethylated	17	44

Figure 7

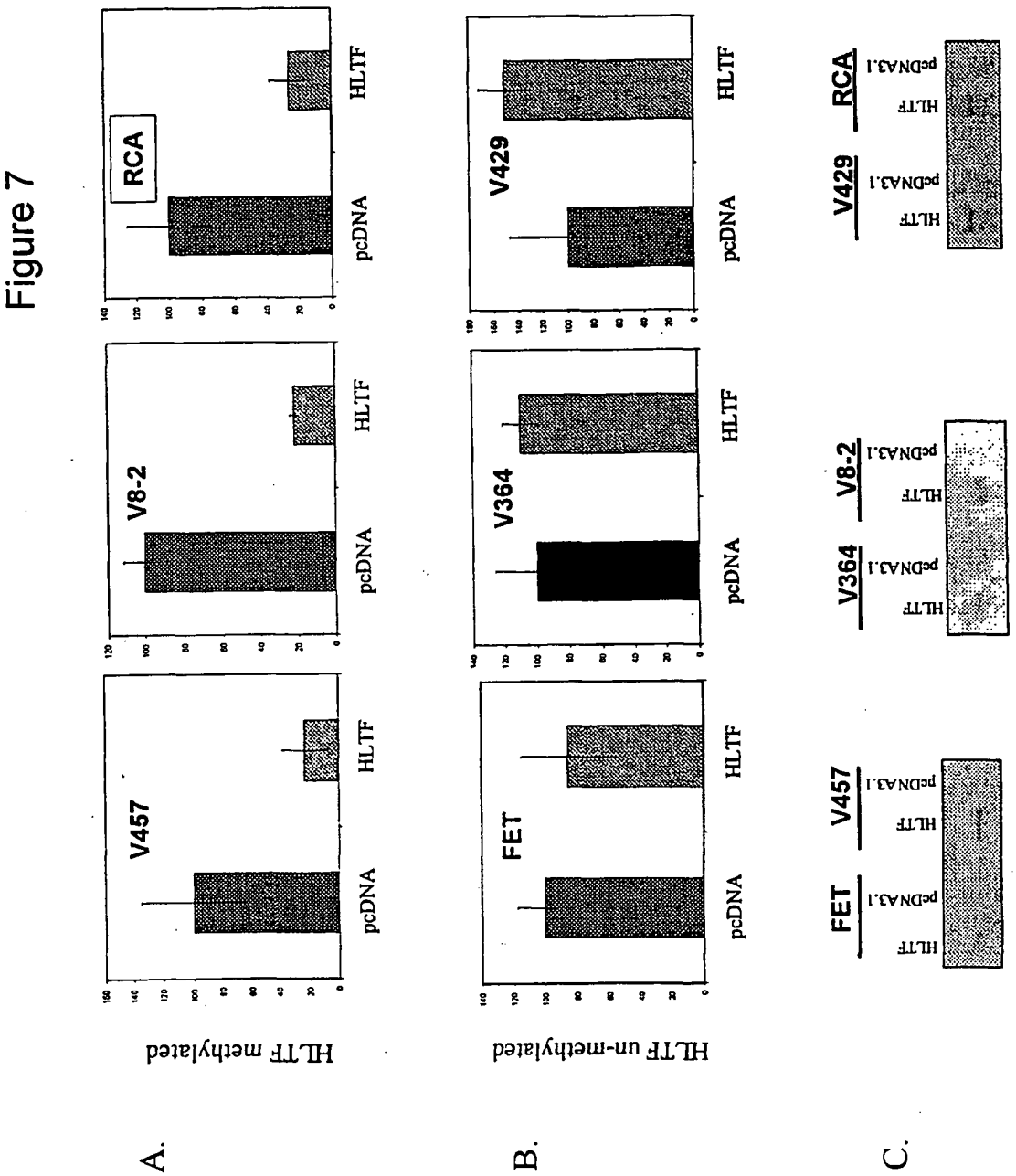
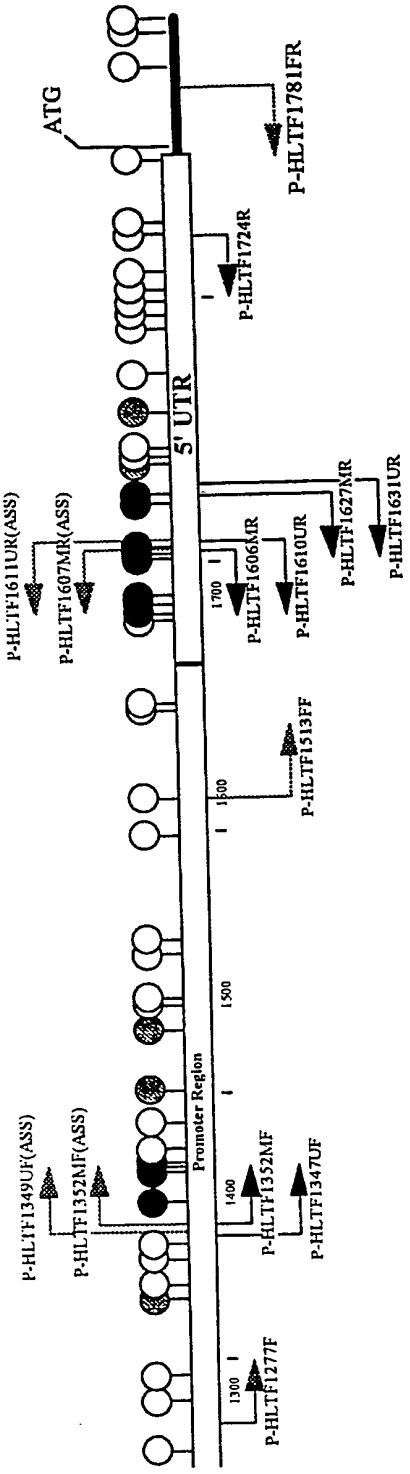

A.
HLTF methylated
V457
RCA
V8-2
pcDNA
HLTF
B.
HLTF un-methylated
FET
V364
V429
C.
FET
V457
V364
V8-2
V429
RCA
HLTF
pcDNA3.1

Figure 8



Primer set (1277F + 1724R): Hpa II assay.
Primer sets: (1352MF +1606MR), (1347UF + 1610UR): MS-PCR assay.
(1352MF +1627MR), (1347UF + 1631UR): MS-PCR assay.
(1352MF(ASS) + 1607MR(ASS), (1349UF(ASS) + 1611UR(ASS)): MS-PCR assay.

Figure 9

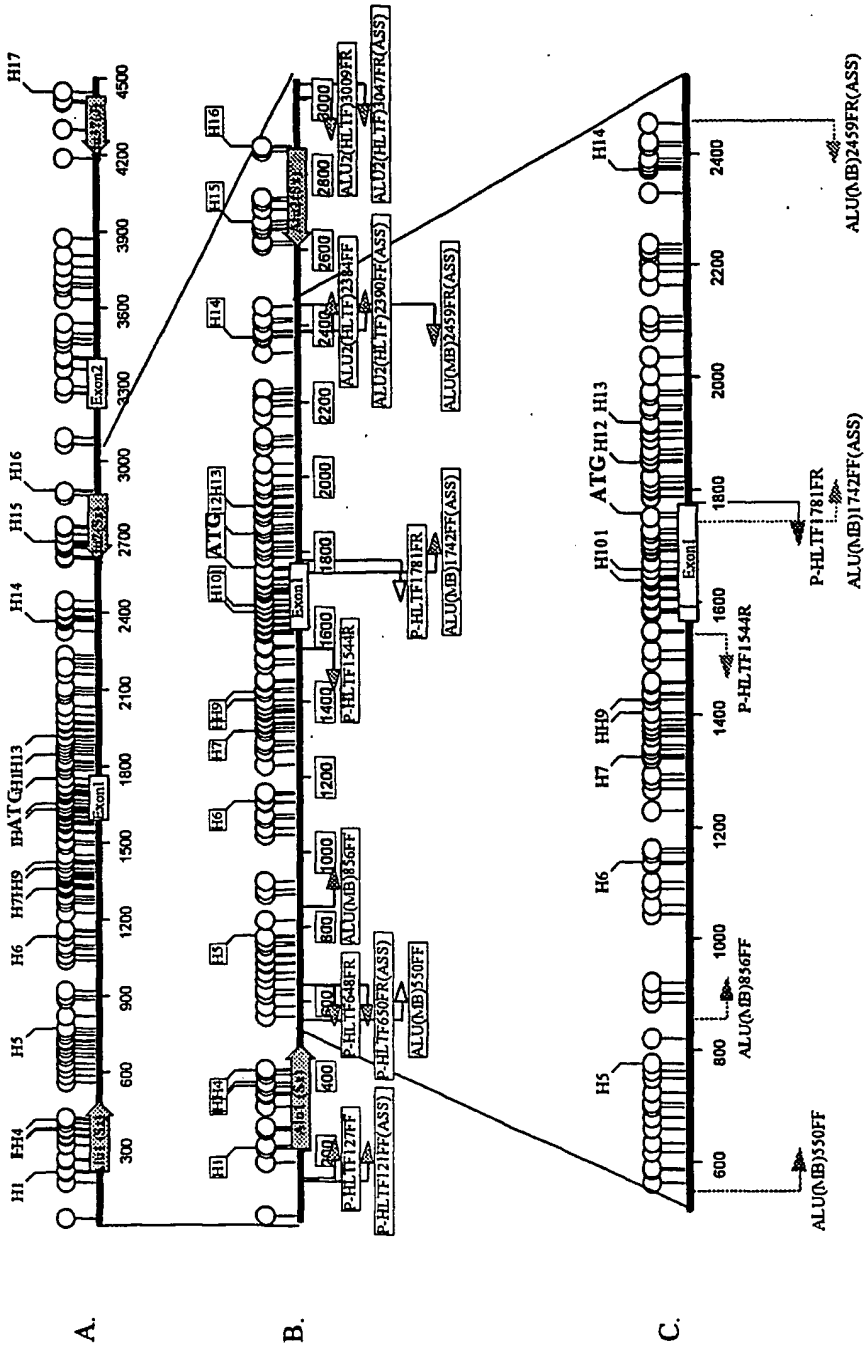

A.
B.
C.
ATG
Exon1
Exon2
P-HLTF127FF
P-HLTF1211FF(ASS)
P-HLTF648FR
P-HLTF650FR(ASS)
ALU(MB)836FF
ALU(MB)550FF
P-HLTF1544R
P-HLTF1781FR
ALU(MB)1742FF(ASS)
ALU(MB)2459FR(ASS)
ALU2(HLTF)2304FF
ALU2(HLTF)2390FF(ASS)
ALU2(HLTF)3009FR
ALU2(HLTF)3047FR(ASS)

Figure 10

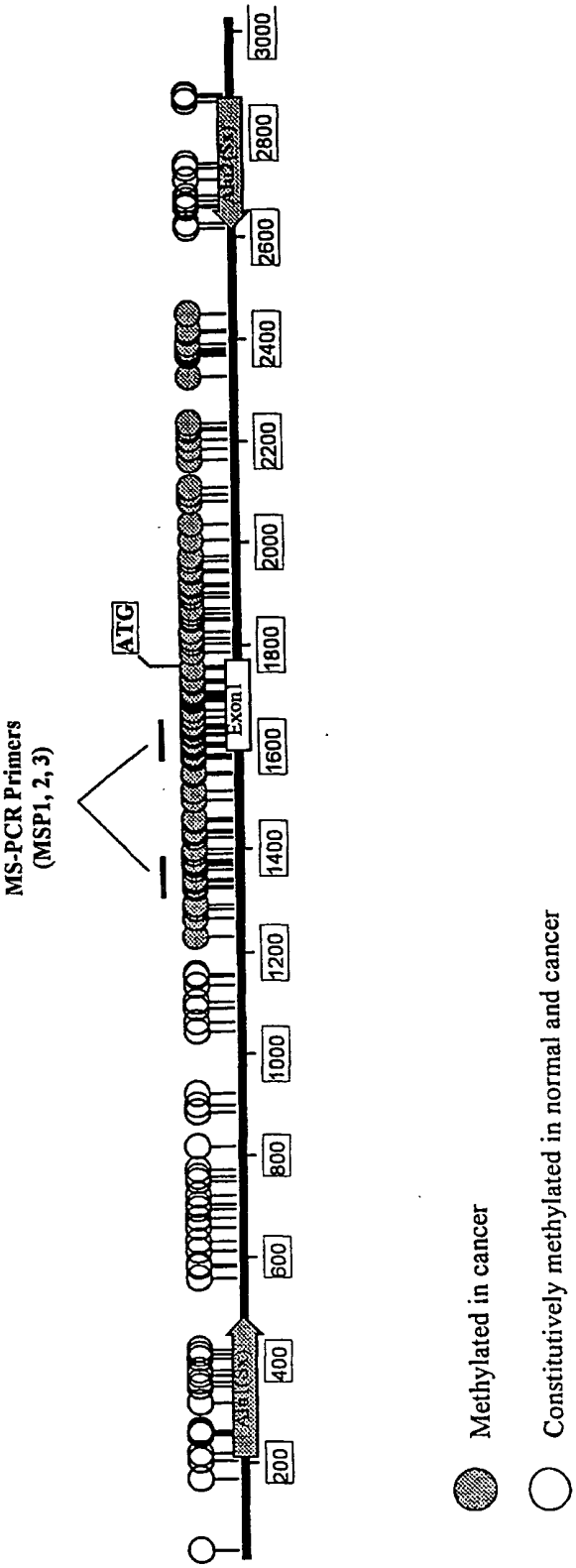

MS-PCR Primers
(MSP1, 2, 3)
ATG
Exon1
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
Methylated in cancer
Constitutively methylated in normal and cancer

Figure 11

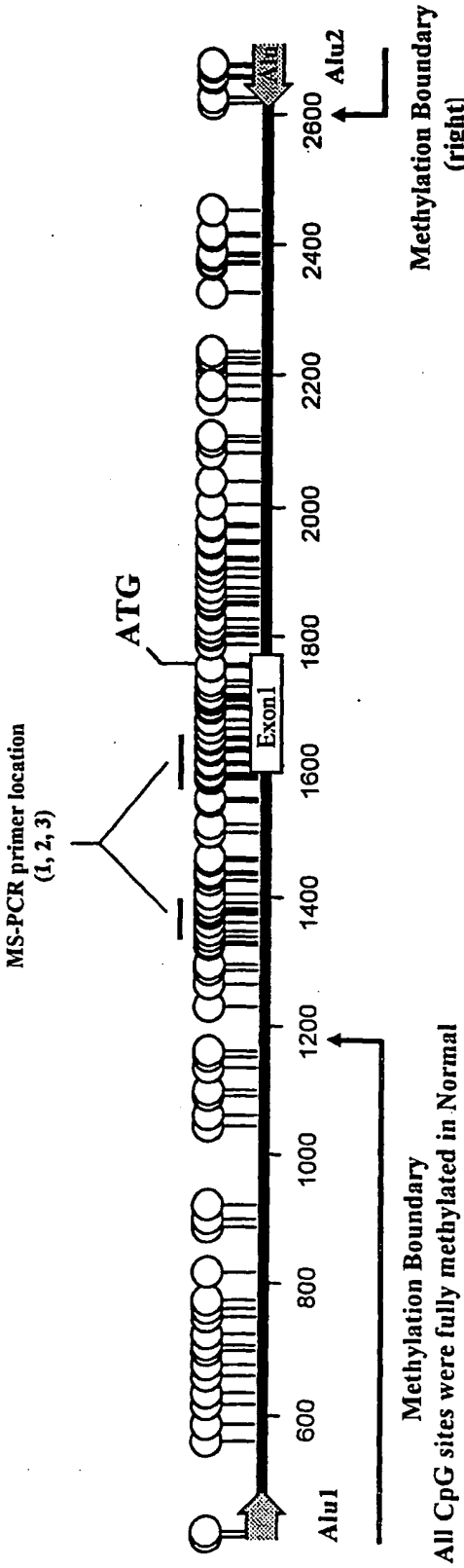
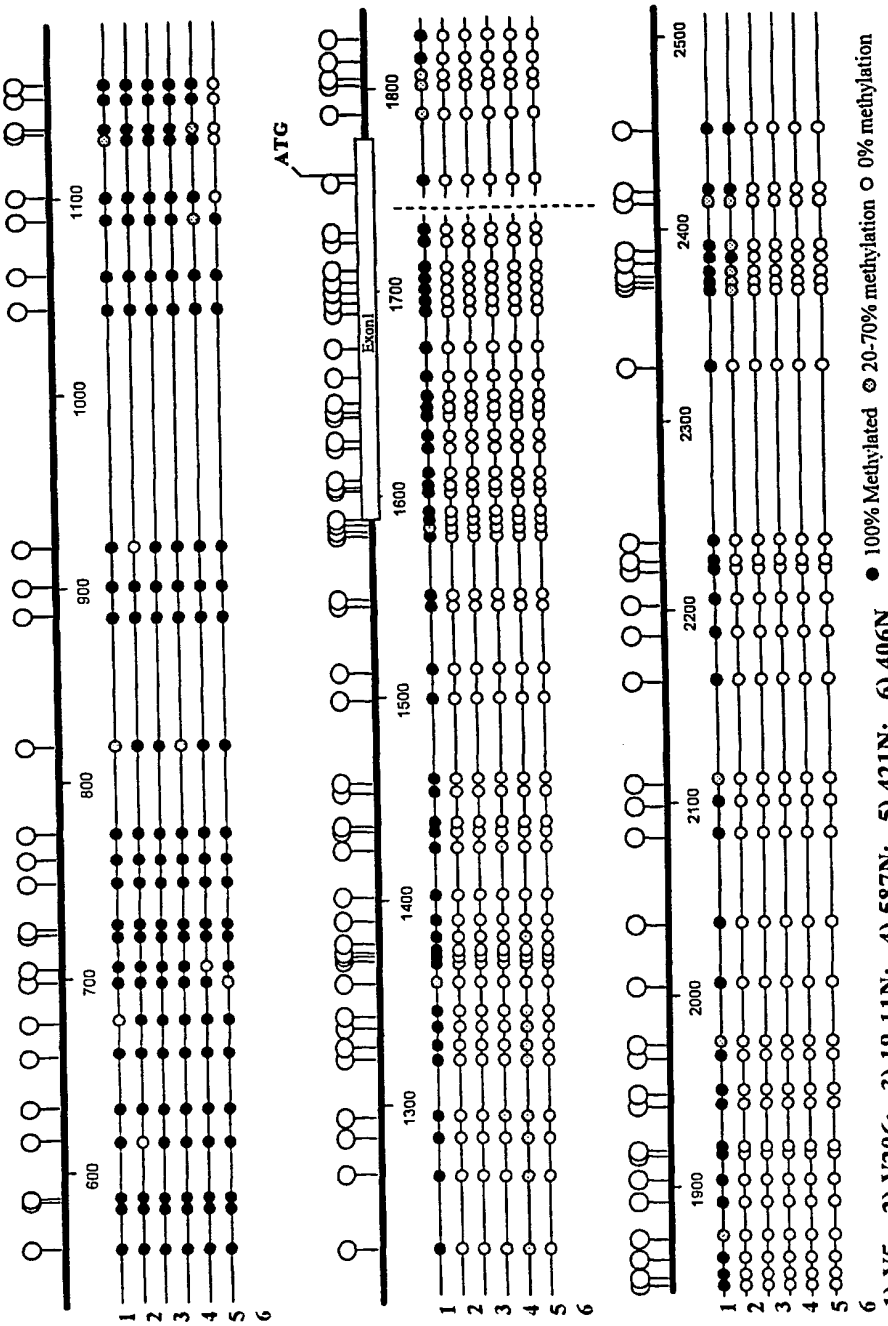

MS-PCR primer location
(1, 2, 3)
ATG
Exon1
Alu1
Alu2
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
Methylation Boundary
All CpG sites were fully methylated in Normal
(left)
Methylation Boundary
(right)

Figure 12



1) V5; 2) V206; 3) 19-11N; 4) 587N; 5) 421N; 6) 406N. ● 100% Methylated ⊗ 20-70% methylation ○ 0% methylation

Figure 13

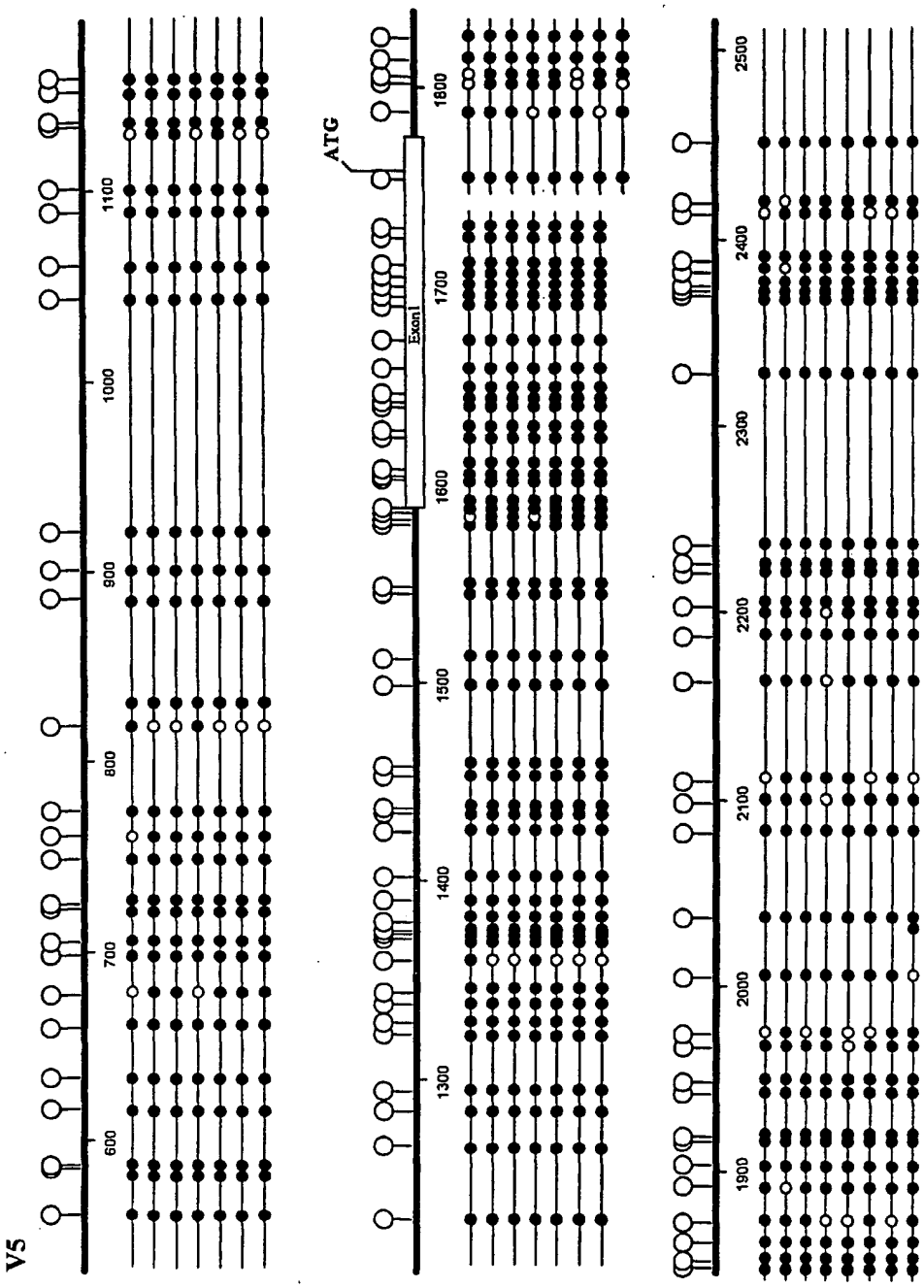

V5
600
700
800
900
1000
1100
1300
1400
1500
1600
ATG
Exon1
1700
1800
1900
2000
2100
2200
2300
2400
2500

Figure 14

mswmfkrdpvwkylqtvqygvhgnfprlsyptffprfefqdvippddfltsdeevdsvlfgslrghvvglryytgvvnnn
emvalqrdpnnpydknaikvnnvngnqvghlkkelagalayimdnklaqiegvvpfgannaftmplhmtfwgkeenrkav
sdqlkkhgfklgpapktlgfnlesgwgsgragpsysmpvhaavqmtteqlktefdklfedlkeddkthemepaeaietpl
lphqkqalawmvsrenskelppfweqrndlyyntitnfsekdrpenvhggiladdmglgktltaiaviltnfhdgrplpi
ervkknllkkeynvnddsmklggnntsekadglskdasrcseqpsisdikekskfrmselsssrpkrrktavqyiessds
eeietselpqkmkgklknvqsetkgrakagsskviedvafacaltssvpttkkkmlkkgacavegskktdveerprttli
icplsvlsnwidqfgqhiksdvhlnfyvyygpdrirepallskqdivlttynilthdygtkgdsplhsirwlrvildegh
airnpnaqqtkavldleserrwvltgtpiqnslkdlwsllsflklkpfidrewwhrtiqrpvtmgdegglrrlqslikni
tlrrtktskikgkpvlelperkvfiqhitlsdeerkiyqsvknegratigryfnegtvlahyadvlglllrlrqicchty
lltnavssngpsgndtpeelrkklirkmkllilssgsdeecaicldsltvpvithcahvfckpcicqviqneqphakcplc
rndihednllecppeelardsekksdmewtssskinalmhaltdlrkknpnikslvvsqfttflslieiplkasgfvftr
ldgsmaqkkrvesiqcfqnteagsptimllslkaggvglnlsaasrvflmdpawnpaaedqcfdrchrlgqkqeviitkf
ivkdsveenmlkiqnkkrelaagafgtkkpnademkqakineirtlidl

Figure 15

5'-

ATTGCCATTTGCCTCAACACACTTCAAGAAGAACATACAACAATTAAAATAATTTGAAATGCCATTGTCTTTATCTTCCTCATCAACTTTGGTCAAATGAAATAATACATT
GTTTCAAATTCTATTGCTCATCCATGGTTCATGATTAGAGATCACTACAGTTTATTTACGATAATGAACAAATGTGTCTTGAAAAGCTGACTACCGGCCTGGCACGGTGGC
TCATGCCTGTAATCCCAGCACTCTGGGAGGCCGAGGCGGGCGGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAATATGGTGAAACCCCGTCTCTACTAAAAA
TACAAAAATTAGCTGGGCGTGGTGGCGGGCTCCTGCAATCCCGGCTACTCCGGAGGCTGAGGCAGGAGAATCGCTTGAATCCGGGAGGCGGAGGTTGCAGTGAGCCAA
GATCACCCCCATTGCACTCCAGCCTGGGGGACAGAGCAAGACTCCATCTCAATAAAAAAGAAAAAAAAACAAAAAAAAGAAAAAAGAAAAGCTGACCACCTAATATTA
TTGTGCTGACAGGCGTGTAACATGGACCTCTCTGTGACGCGCAAGTATTATCTGGCAGTTACATGGTGCCGCCTGTGCAGCTGGGCGCCACCACCAGCTTCCTGCCCTGT
GCGATTAAAAGGCCAGCTTCGGGGACAGCAGCCTCCTCAACGCAAAGCGACCAGCCCTGCCAGGCGTCGGCTGCAGCTGGTCCAAGGTTGACGGTCCCAGGGCCGCATT
CTCCTTCCGGTCACCTCAGACTCTGGAAAGAGAGCCATGGAGGGCCTCAGAGCGCCCCTCTGCAGTTCCTTGAGGAAAACAAGGGCCTCCCCCTCTTGGCTGGAGCAGCC
AAGTGCCCACGGGGAAAATGGTCCCGTCCACCATTTCAGGCAAACGCCCCCATAGCTGTAATAAACATGTGATGGGTTTCCTGCACCAAGAAAGGAAAATTTTCCTTAA
AAATAACAAAAAGAGGTGGGGAGGGGACTCTAAGTTCCATTTAAGATTTTCCAAACTCGTTGTGTCTCATGCATACGCTGAGGCTTAACATGAAAATTTGGAGCGAAAG
TCCCCACGGTTCACCACAGGGAATGAAGTAAAGTATCTCGCCGGTAGCTCTAGGTCTCGTAAACCGTGGAGCTCCTTTAGTATAAGGCCCTATTATGTGCCACATTAGAT
GAGAAGAAAATAAGAGGAAGGATACGGTAGAGCCTTTAGAATGGGGTTCAGTCTTAGGACTCGTAGGATAAAGGAAGGTCGTTTCCCTCCGTTTGAGGCAGGGAATCA
AACAAAACACCGGCACCGCAGGCACCGCAGTCGCACTCCTGGGGCCTCGTGGCTTTCCCGCGCGCCCGCCTTGGGGGCGGGGAAGAACCCGGATGGAACCACCCTTGCA
GCCCGGACCCCCCCGCCGTCTTTGAATGGCAGCGGGGCGGAGTTGCCCTCCCTTCTGTGCTCTGACTGGTTTGGCTCCGCACTCTCCTCTTCGTGATTGGGCTTTCCTAGTGC
CAGTCACAGAGCGACGCTGGTCTCCCAGATTGTTGCAGAAGGAGACGGCGTCGACGTCTGACTGGACTCGCGGCGACTTACCTTTCAGTCGTGCGCTCCTGATCCGGCGC
TCGGAATTTGTCCCCGGCTTCAGGGCTGCGGGGCCTGGAAGGAGGCGTATCGAGGCGGCTCGAAAACGATCCAGGGGAGCCGAGGCGCTCCTCTTGTCATCCCACTCAG
CGCCATGTCCTGGATGTTCAAGAGGTGAAGGGGGCGGAGGGGGTGGGGCGCTCGGTCTAACGGCCTGGAGGCGTCCCCAAATGAACCTGACCTTCCCGGCGTTCCTCT
GCGTTCCCTGGGCGATTTGTGCAGCTGTATTCGTTCTGTTGGTCGCATATGTGGCCGCCGGAGAAATAAATGCATTGTCTTCGCTGGCGAGTAGGGGCTCCTAGGGCGAG
TCCCGTGTTAGGGACTTGGGAAATCTCTGTCACCGACTGTGGGGGGCCTGGGGCTTAAAGCATCTCGGCCAGTGCTTTATTTAACAAAACATCCATTAGATACCCACTGT
GCGCCAGGCATTGTGCTCGGTGGAGGGCCGAGTAAGACAGTCCTTGTCCCCAAGTGGCTCACAGCTTATCTGTATGCACAGACGTGCAGATAAGTATTTGCAGAGGCGCT
AGTAGAGATATGCGCCACTTGCAAGGAAAACGCATCGCAAATAAGCGACTGTTCTTAGGAGAATGGAAACTTTTTGCAATGCTCCAGTAGACTCCCCTTCACTTTTCATT
GTCCAAATTGGTATTTATGGTCATTCCCGAAACAGTTACACCTAATGGGTTTGGGGTTATGCTTAAACCGGCGACGGAAGGCGGGTACGGGTGGCTTGGGGCCAATTTCC
CACGGTGGCGTGGTTTGGATACCAGAATAGGTGTTCTGCTCGCAGGGAGTTACAGTGATGGATGCTGGATGTATTTTCTGGAAAATTATAAACAACTTCTGGTACATTAA
TAGTTGTATTTTTATAAGTCTTTTATATAGCTCAATCCCTATTGAAATTTTGGCATATGATTTCATTTTTTAAATTATTTTTTTGACATGGAGTCTCGCTCTGTCGTCCAGGC
TGGAGTGCAGTGGCACAATCTCGGCTCGCTGCAACCTACGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACATATGGGCGCCACCAC
GCCTGGCTAATTTTTGTATTTGTAGTAGACAGGGTTTCACAATGTTGGCCAGGTTGGTCTCAAATTCCTGACCTCAGGTGATCCACCTGCCTCTGCCTCCCAAAGTGCTGG
GATTCCAGGCGTGAGCCACCGCATCCGGCCCATAATTGTAGTTACAGACATCATATACATATCAATTTAGATGATTTTCACTGGGAAAGAGGTTAGTGAACTCTGACTAG
GCTTTTCATTTTGAATACCTAGGGGGCAGGTAGAAG -3'

Figure 16

5'-
TGGCAGTTACATGGTGCCGCCTGTGCAGCTGGGCGCCACCACCAGCTTCCTGCCCTGTGCGATTAAAAGGCCAGCTTCGG
GGACAGCAGCCTCCTCAACGCAAAGCGACCAGCCCTGCCAGGCGTCGGCTGCAGCTGGTCCAAGGTTGACGGTCCCAGGG
CCGCATTCTCCTTCCGGTCACCTCAGACTCTGGAAAGAGAGCCATGGAGGGCCTCAGAGCGCCCCTCTGCAGTTCCTTGA
GGAAAACAAGGGCCTCCCCCTCTTGGCTGGAGCAGCCAAGTGCCCACGGGGAAAATGGTCCCGTCCACCATTTCAGGCAA
ACGCCCCCATAGCTGTAATAAACATGTGATGGGTTTCCTGCACCAAGAAAGGAAAATTTTCCTTAAAAATAACAAAAAGA
GGTGGGGAGGGGACTCTAAGTTCCATTTAAGATTTTCCAAACTCGTTGTGTCTCATGCATACGCTGAGGCTTAACATGAA
AATTTGGAGCGAAAGTCCCCACGGTTCACCACAGGGAATGAAGTAAAGTATCTCGCCGGTAGCTCTAGGTCTCGTAAACC
GTGGAGCTCCTTTAGTATAAGGCCCTATTATGTGCCACATTAGATGAGAAGAAAATAAGAGGAAGGATACGGTAGAGCCT
TTAGAATGGGGTTCAGTCTTAGGACTCGTAGGATAAAGGAAGGTCGTTTCCCTCCGTTTGAGGCAGGGAATCAAACAAAA
CACCGGCACCGCAGGCACCGCAGTCGCACTCCTGGGGCCTCGTGGCTTTCCCGCGCGCCCGCCTTGGGGGCGGGGAAGAA
CCCGGATGGAACCACCCTTGCAGCCCGGACCCCCCGCCGTCTTTGAATGGCAGCGGGGCGGAGTTGCCCTCCCTTCTGTG
CTCTGACTGGTTTGGCTCCGCACTCTCCTCTTCGTGATTGGGCTTTCCTAGTGCCAGTCACAGAGCGACGCTGGTCTCCC
AGATTGTTGCAGAAGGAGACGGCGTCGACGTCTGACTGGACTCGCGGCGACTTACCTTTCAGTCGTGCGCTCCTGATCCG
GCGCTCGGAATTTGTCCCCGGCTTCAGGGCTGCGGGGCCTGGAAGGAGGCGTATCGAGGCGGCTCGAAAACGATCCAGGG
GAGCCGAGGCGCTCCTCTTGTCATCCCACTCAGCGCCATGTCCTGGATGTTCAAGAGGTGAAGGGGGCGGAGGGGGTGGG
GCGCTCGGTCTAACGGCCTGGAGGCGTCCCCAAATGAACCTGACCTTCCCGGCGTTCCTCTGCGTTCCCTGGGCGATTTG
TGCAGCTGTATTCGTTCTGTTGGTCGCATATGTGGCCGCCGGAGAAATAAATGCATTGTCTTCGCTGGCGAGTAGGGGCT
CCTAGGGCGAGTCCCGTGTTAGGGACTTGGGAAATCTCTGTCACCGACTGTGGGGGGCCTGGGGCTTAAAGCATCTCGGC
CAGTGCTTTATTTAACAAAACATCCATTAGATACCCACTGTGCGCCAGGCATTGTGCTCGGTGGAGGGCCGAGTAAGACA
GTCCTTGTCCCCAAGTGGCTCACAGCTTATCTGTATGCACAGACGTGCAGATAAGTATTTGCAGAGGCGCTAGTAGAGAT
ATGCGCCACTTGCAAGGAAAACGCATCGCAAATAAGCGACTGTTCTTAGGAGAATGGAAACTTTTTGCAATGCTCCAGTA
GACTCCCCTTCACTTTTCATTGTCCAAATTGGTATTTATGGTCATTCCCGAAACAGTTACACCTAATGGGTTTGGGGTTA
TGCTTAAACCGGCGACGGAAGGCGGGTACGGGTGGCTTGGGGCCAATTTCCCACGGTGGCGTGGTTTGGATACCAGAATA
GGTGTTCTGCTCGCAGGGAGTTACAGTGATGGATGCTGGATGTATTTTCTGGAAAATTATAAACAACTTCTGGTACATTA
ATAGTTGTATTTTTATAAGTCTTTTATATAGCTCAATCCCTATTGAAATTTTGGCATATGATTTCATTTTTTAAATTATT
T-3'

Figure 17

5'-
GTTCAGTCTTAGGACTCGTAGGATAAAGGAAGGTCGTTTCCCTCCGTTTGAGGCAGGGAATCAAACAAAACACCGGCACCGCAGGCACCG

CAGTCGCACTCCTGGGGCCTCGTGGCTTTCCCGCGCGCCCGCCTTGGGGGCGGGAAGAACCCGGATGGAACCACCCTTGCAGCCCGGAC

CCCCCGCCGTCTTTGAATGGCAGCGGGGCGGAGTTGCCCTCCCTTCTGTGCTCTGACTGGTTTGGCTCCGCACTCTCCTCTTCGTGATTGGG

CTTTCCTAGTGCCAGTCACAGAGCGACGGCTGGTCTCCCAGATTGTTGCAGAAGGAGACGGCGTCGACGTCTGACTGGACTCGCCGGCGACT

TACCTTTCAGTCGTGCGCTCCTGATCCGGCTCGGAATTTGTCCCCGGCTTCAGGGCTGCGGGGCCTGGAAGGAGGCGTATCGAGGCGG

CTCGAAAACGATCCAGGGGAGCCGAGGCGCTCCTCTTGTCATCCCACTCAGCGCCATGTCCTGGATGTTCAAGAGGTGAAGGGGGCGGA

GGGGGTGGGG-3'

Figure 18

Sense Strand-Methylated (Bisulfite conversion):

5'-GTTTAGTTTTAGGATTCGTAGGATAAAGGAAGGTCGTTTTTTTTCGTTTGAGGTAGGGAATTAAATAAAATATCGGTATCGTAGGTATCG

TAGTCGTATTTTTGGGGTTTCGTGGTTTTTTTCGCGCGTTCGTTTTGGGGGCGGGGAAGAATTCGGATGGAATTATTTTTGTAGTTCGGAT

TTTTCGTCGTTTTTGAATGGTAGCGGGGCGGAGTTGTTTTTTTTTTTGTGTTTTGATTGGTTTGGTTTCGTATTTTTTTTTTTCGTGATTGGG

TTTTTTTAGTGTTAGTTATAGAGCGACGTTGGTTTTTTAGATTGTTGTAGAAGGAGACGGCGTCGACGTTTGATTGGATTCGCGGCGATT

TATTTTTTAGTCGTGCGTTTTTGATTCGGCGTTCGGAATTTGTTTTCGGTTTTAGGGTTGCGGGGTTTGGAAGGAGGCGTATCGAGGCGG

TTCGAAAACGATTTAGGGGAGTCGAGGCGTTTTTTTTGTTATTTTATTTAGCGTTATGTTTTGGATGTTTAAGAGGTGAAGGGGGCGGA

GGGGGTGGGG-3'

Sense Strand-Unmethylated (Bisulfite conversion):

5'-GTTTAGTTTTAGGATTTGTAGGATAAAGGAAGGTTGTTTTTTTTTGTTTGAGGTAGGGAATTAAATAAAATATTGGTATTGTAGGTATTG

TAGTTGTATTTTTGGGGTTTTGTGGTTTTTTTGTGTGTTTGTTTTGGGGGTGGGGAAGAATTTGGATGGAATTATTTTTGTAGTTTGGAT

TTTTTGTTGTTTTTGAATGGTAGTGGGGTGGAGTTGTTTTTTTTTTTGTGTTTTGATTGGTTTGGTTTTGTATTTTTTTTTTTTGTGATTGGG

TTTTTTTAGTGTTAGTTATAGAGTGATGTTGGTTTTTTAGATTGTTGTAGAAGGAGATGGTGTTGATGTTTGATTGGATTTGTGGTGATT

TATTTTTTAGTTGTGTGTTTTTGATTTGGTGTTTGGAATTTGTTTTTGGTTTTAGGGTTGTGGGGTTTGGAAGGAGGTGTATTGAGGTGG

TTTGAAAATGATTTAGGGGAGTTGAGGTGTTTTTTTTGTTATTTTATTTAGTGTTATGTTTTGGATGTTTAAGAGGTGAAGGGGGTGGA

GGGGGTGGGG-3'

Figure 19

Complementary Strand-Methylated (Bisulfite conversion):

3'-CAAATCAAAATCCTAAGCATCCTATTTCCTTCCAGCAAAAAAAAGCAAACTCCATCCCTTAATTTATTTTATAGCCATAGCATCCATAGC

ATCAGCATAAAAACCCCAAAGCACCAAAAAAGCGCGCAAGCAAAACCCCCGCCCCTTCTTAAGCCTACCTTAATAAAAACATCAAGCCTA

AAAAGCAGCAAAAACTTACCATCGCCCCGCCTCAACAAAAAAAAAAACACAAAACTAACCAAACCAAAGCATAAAAAAAAAAGCACTAACCC

AAAAAAATCACAATCAATATCTCGCTGCAACCAAAAAATCTAACAACATCTTCCTCTGCCGCAGCTGCAAACTAACCTAAGCGCCGCTAA

ATAAAAAATCAGCACGCAAAAACTAAGCCGCAAGCCTTAAACAAAAGCCAAAATCCCAACGCCCCAAACCTTCCTCCGCATAGCTCCGCC

AAGCTTTTGCTAAATCCCCTCAGCTCCGCAAAAAAAACAATAAAATAAATCGCAATACAAAACCTACAAATTCTCCACTTCCCCCGCCT

CCCCCACCCC-5'

Complementary Strand-Unmethylated (Bisulfite conversion):

3'-CAAATCAAAATCCTAAACATCCTATTTCCTTCCAACAAAAAAAAACAAACTCCATCCCTTAATTTATTTTATAACCATAACATCCATAAC

ATCAACATAAAAACCCCAAAACACCAAAAAAAACACACACAAACAAAACCCCCACCCCTTCTTAAACCTACCTTAATAAAAACATCAAACCTA

AAAAACAACAAAAACTTACCATCACCCCACCTCAACAAAAAAAAAAACACAAAACTAACCAAACCAAAACATAAAAAAAAAAACACTAACCC

AAAAAAATCACAATCAATATCTCACTACAACCAAAAAAATCTAACAACATCTTCCTCTACCACAACTACAAACTAACCTAAACACCACTAA

ATAAAAAATCAACACACAAAAACTAACCACAAACCTTAAACAAAAACCAAAATCCCAACACCCCAAACCTTCCTCCACATAACTCCACC

AAACTTTTACTAAATCCCCTCAACTCCACAAAAAAAACAATAAAATAAATCACAATACAAAACCTACAAATTCTCCACTTCCCCCACCT

CCCCCACCCC-5'

Figure 20

A.	HpaII assay	P-HLTF1277F (HpaII)	5'-GGAAGGTCGTTTCCCTCCGTTTGAG -3'
		P-HLTF1724R (HpaII)	5'-CGGCTCCCCTGGATCGTTTTCGAG -3'
B.	MSP1 (Sense Strand)	P-HLTF1352MF(MSP-1&2)	5'-TGGGGTTTCGTGGTTTTTTCGCGC-3'
		P-HLTF1606MR (MSP-1)	5'-CCGCGAATCCAATCAAACGTCGACG-3'
		P-HLTF1347UF (MSP-1&2)	5'-ATTTTTGGGGTTTTGTGGTTTTTTTGTGT-3'
		P-HLTF1610UR (MSP-1)	5'-ATCACCACAAATCCAATCAAACATCAACA-3'
C.	MSP2 (Sense Strand)	P-HLTF1352MF(MSP-1&2)	5'-TGGGGTTTCGTGGTTTTTTCGCGC-3'
		P-HLTF1627MR (MSP-2)	5'-GCACGACTAAAAAATAAATCGCCGCG-3'
		P-HLTF1347UF (MSP-1&2)	5'-ATTTTTGGGGTTTTGTGGTTTTTTTGTGT-3'
		P-HLTF1631UR (MSP-2)	5'-AAACACACAACTAAAAAATAAATCACCACA-3'
D.	MSP3 (Anti-Sense Strand)	P-HLTF1352MF(ASS) (MSP-3)	5'-TAAAACCTCGTAACTTTCCCGCGCG-3'
		P-HLTF1607MR(ASS) (MSP-3)	5'-GTCGCGAGTTTAGTTAGACGTCGAC-3'
		P-HLTF1349UF(ASS) (MSP-3)	5'-TCCTAAAACCTCATAACTTTCCCACACA-3'
		P-HLTF1611UR(ASS) (MSP-3)	5'-AGTTGTTGTGAGTTTAGTTAGATGTTGAT-3'

Figure 21

5'-
TAGATGAGAAGAAAATAAGAGGAAGGATACGGTAGAGCCTTTAGAATGGGGTTCAGTCTTAGGACTCGTAGGATAAAGGA
AGGTCGTTTCCCTCCGTTTGAGGCAGGGAATCAAACAAAACACCGGCACCGCAGGCACCGCAGTCGCACTCCTGGGGCCT
CGTGGCTTTCCCGCGCCCGCCTTGGGGCGGGGAAGAACCCGGATGGAACCACCCTTGCAGCCCGGACCCCCCGCCGT
CTTTGAATGGCAGCGGGGCGGAGTTGCCCTCCCTTCTGTGCTCTGACTGGTTTGGCTCCGCACTCTCCTCTTCGTGATTG
GGCTTTCCTAGTGCCAGTCACAGAGCGACGCTGGTCTCCCAGATTGTTGCAGAAGGAGACGGCGTCGACGTCTGACTGGA
CTCGCGGCCGACTTACCTTTCAGTCGTGCGCTCCTGATCCGGCGCTCGGAATTTGTCCCCGGCTTCAGGGCTGCGGGGCCT
GGAAGGAGGCGTATCGAGGCGGCTCGAAAACGATCCAGGGGAGCCGAGGCGCTCCTCTTGTCATCCCACTCAGCGCCATG
TCCTGGATGTTCAAGAGGTGAAGGGGGCGGAGGGGGTGGGGCGCTCGGTCTAACGGCCTGGAGGCGTCCCCAAATGAACC
TGACCTTCCCGGCGTTCCTCTGCGTTCCCTGGGCGATTTGTGCAGCTGTATTCGTTCTGTTGGTCGCATATGTGGCCGCC
GGAGAAATAAATGCATTGTCTTCGCTGGCGAGTAGGGGCTCCTAGGGCGAGTCCCGTGTTAGGGACTTGGGAAATCTCTG
TCACCGACTGTGGGGGGCCTGGGGCTTAAAGCATCTCGGCCAGTGCTTTATTTAACAAAACATCCATTAGATACCCACTG
TGCGCCAGGCATTGTGCTCGGTGGAGGGCCGAGTAAGACAGTCCTTGTCCCCAAGTGGCTCACAGCTTATCTGTATGCAC
AGACGTGCAGATAAGTATTTGCAGAGGCGCTAGTAGAGATATGCGCCACTTGCAAGGAAAACGCATCGCAAATAAGCGAC
TGTTCTTAGGAGAATGGAAACTTTTTGCAATGCTCCAGTAGACTCCCCTTCACTTTTCATTGTCCAAATTGGTATTTATG
GTCATTCCCGAAACAGTTACACCTAATGGGTTTGGGGTTATGCTTAAACCGGCGACGGAAGGCGGGTACGGGTGGCTTGG
GGCCAATTTCCCACGGTGGCGTGGTTTGGATACCAGAATAGGTGTTCTGCTCGCAGGGAGTTACAGTGATGGATGCTGGA
TGTATTTTCTGGAAAATTATA-3'

Figure 22

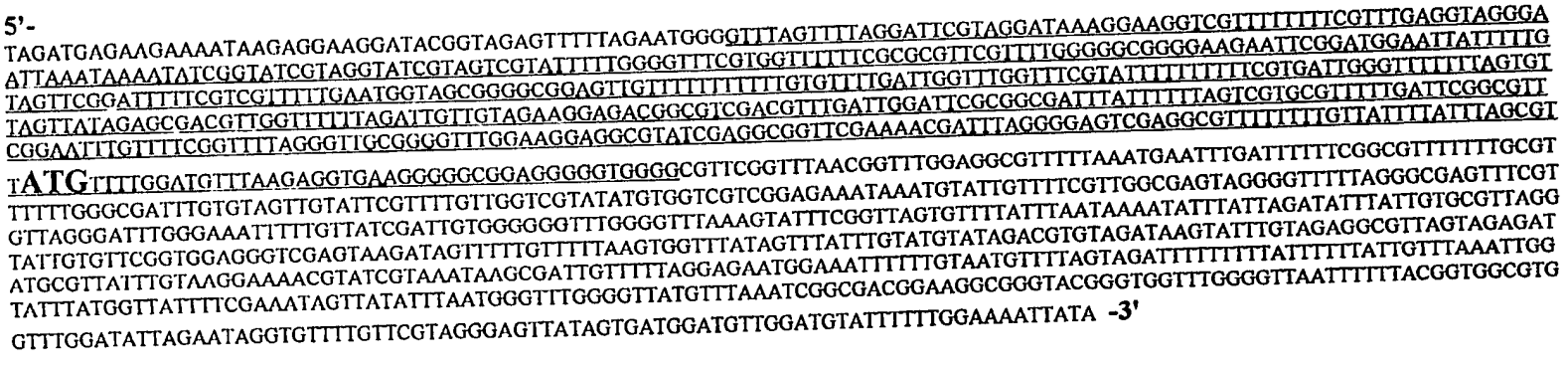

5'-
TAGATGAGAAGAAAATAAGAGGAAGGATACGGTAGAGTTTTTAGAATGGGGTTTAGTTTTAGGATTCGTAGGATAAAGGAAGGTCGTTTTTTTTCGTTTGAGGTAGGGA
ATTAAATAAAATATCGGTATCGTAGGTATCGTAGTCGTATTTTTGGGGTTTCGTGGTTTTTTCGCGCGTTCGTTTTGGGGGCGGGGAAGAATTCGGATGGAATTATTTTTG
TAGTTCGGATTTTTCGTCGTTTTTGAATGGTAGCGGGGCGGAGTTGTTTTTTTTTTTGTGTTTTGATTGGTTTGGTTTCGTATTTTTTTTTTCGTGATTGGGTTTTTTTAGTGT
TAGTTATAGAGCGACGTTGGTTTTTTAGATTGTTGTAGAAGGAGACGGCGTCGACGTTTGATTGGATTCGCGGCGATTTATTTTTTAGTCGTGCGTTTTTGATTCGGCGTT
CGGAATTTGTTTTCGGTTTTAGGGTTGCGGGGTTTGGAAGGAGGCGTATCGAGGCGGTTCGAAAACGATTTAGGGGAGTCGAGGCGTTTTTTTTGTTATTTTATTTAGCGT
TATGTTTTGGATGTTTAAGAGGTGAAGGGGGCGGAGGGGGTGGGGCGTTCGGTTTAACGGTTTGGAGGCGTTTTTAAATGAATTTGATTTTTTCGGCGTTTTTTTGCGT
TTTTTGGGCGATTTGTGTAGTTGTATTCGTTTTGTTGGTCGTATATGTGGTCGTCGGAGAAATAAATGTATTGTTTTCGTTGGCGAGTAGGGGTTTTTAGGGCGAGTTTCGT
GTTAGGGATTTGGGAAATTTTTGTTATCGATTGTGGGGGGTTTGGGGTTTAAAGTATTTCGGTTAGTGTTTTATTTAATAAAATATTTATTAGATATTTATTGTGCGTTAGG
TATTGTGTTCGGTGGAGGGTCGAGTAAGATAGTTTTTGTTTTTAAGTGGTTTATAGTTTATTTGTATGTATAGACGTGTAGATAAGTATTTGTAGAGGCGTTAGTAGAGAT
ATGCGTTATTTGTAAGGAAAACGTATCGTAAATAAGCGATTGTTTTTAGGAGAATGGAAATTTTTTGTAATGTTTTAGTAGATTTTTTTTTATTTTTTATTGTTTAAATTGG
TATTTATGGTTATTTTCGAAATAGTTATATTTAATGGGTTTGGGGTTATGTTTAAATCGGCGACGGAAGGCGGGTACGGGTGGTTTGGGGTTAATTTTTTACGGTGGCGTG
GTTTGGATATTAGAATAGGTGTTTTGTTCGTAGGGAGTTATAGTGATGGATGTTGGATGTATTTTTTGGAAAATTATA -3'

Figure 23

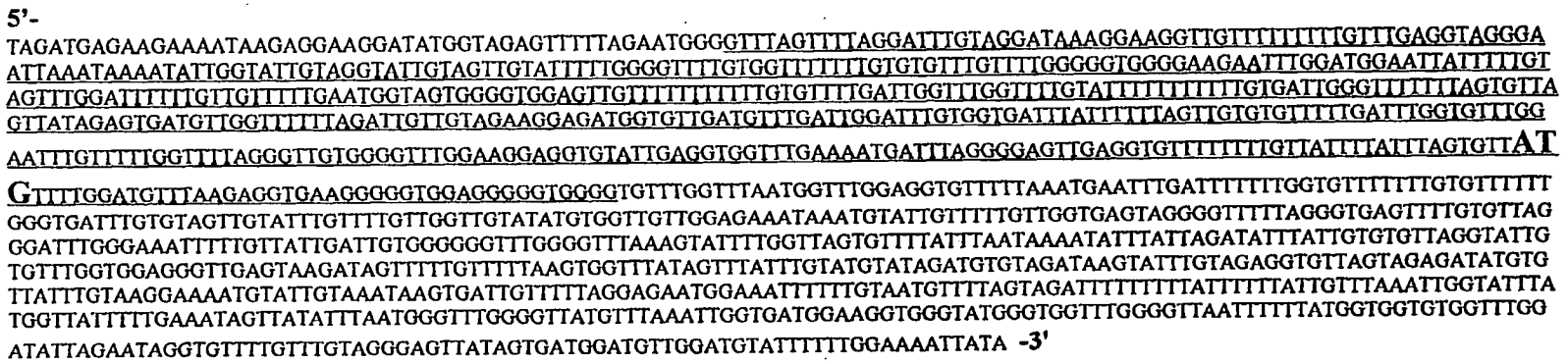

5'-
TAGATGAGAAGAAAATAAGAGGAAGGATATGGTAGAGTTTTTAGAATGGGGTTTAGTTTTAGGATTTGTAGGATAAAGGAAGGTTGTTTTTTTTTGTTTGAGGTAGGGA
ATTAAATAAAATATTGGTATTGTAGGTATTGTAGTTGTATTTTTGGGGTTTTGTGGTTTTTTTGTGTGTTTGTTTTGGGGGTGGGGAAGAATTTGGATGGAATTATTTTTGT
AGTTTGGATTTTTTGTTGTTTTTGAATGGTAGTGGGGTGGAGTTGTTTTTTTTTTTGTGTTTTGATTGGTTTGGTTTTGTATTTTTTTTTTTGTGATTGGGTTTTTTTAGTGTTA
GTTATAGAGTGATGTTGGTTTTTTAGATTGTTGTAGAAGGAGATGGTGTTGATGTTTGATTGGATTTGTGGTGATTTATTTTTTAGTTGTGTGTTTTTGATTTGGTGTTTGG
AATTTGTTTTTGGTTTTAGGGTTGTGGGGTTTGGAAGGAGGTGTATTGAGGTGGTTTGAAAATGATTTAGGGGAGTTGAGGTGTTTTTTTTGTTATTTTATTTAGTGTTAT
GTTTTGGATGTTTAAGAGGTGAAGGGGGTGGAGGGGGTGGGGTGTTTGGTTTAATGGTTTGGAGGTGTTTTTAAATGAATTTGATTTTTTTGGTGTTTTTTTGTGTTTTTT
GGGTGATTTGTGTAGTTGTATTTGTTTTGTTGGTTGTATATGTGGTTGTTGGAGAAATAAATGTATTGTTTTTGTTGGTGAGTAGGGGTTTTTAGGGTGAGTTTTGTGTTAG
GGATTTGGGAAATTTTTGTTATTGATTGTGGGGGGTTTGGGGTTTAAAGTATTTTGGTTAGTGTTTTATTTAATAAAATATTTATTAGATATTTATTGTGTGTTAGGTATTG
TGTTTGGTGGAGGGTTGAGTAAGATAGTTTTTGTTTTTAAGTGGTTTATAGTTTATTTGTATGTATAGATGTGTAGATAAGTATTTGTAGAGGTGTTAGTAGAGATATGTG
TTATTTGTAAGGAAAATGTATTGTAAATAAGTGATTGTTTTTAGGAGAATGGAAATTTTTTGTAATGTTTTAGTAGATTTTTTTTTATTTTTTATTGTTTAAATTGGTATTTA
TGGTTATTTTTGAAATAGTTATATTTAATGGGTTTGGGGTTATGTTTAAATTGGTGATGGAAGGTGGGTATGGGTGGTTTGGGGTTAATTTTTTATGGTGGTGTGGTTTGG
ATATTAGAATAGGTGTTTTGTTTGTAGGGAGTTATAGTGATGGATGTTGGATGTATTTTTTGGAAAATTATA -3'

Figure 24

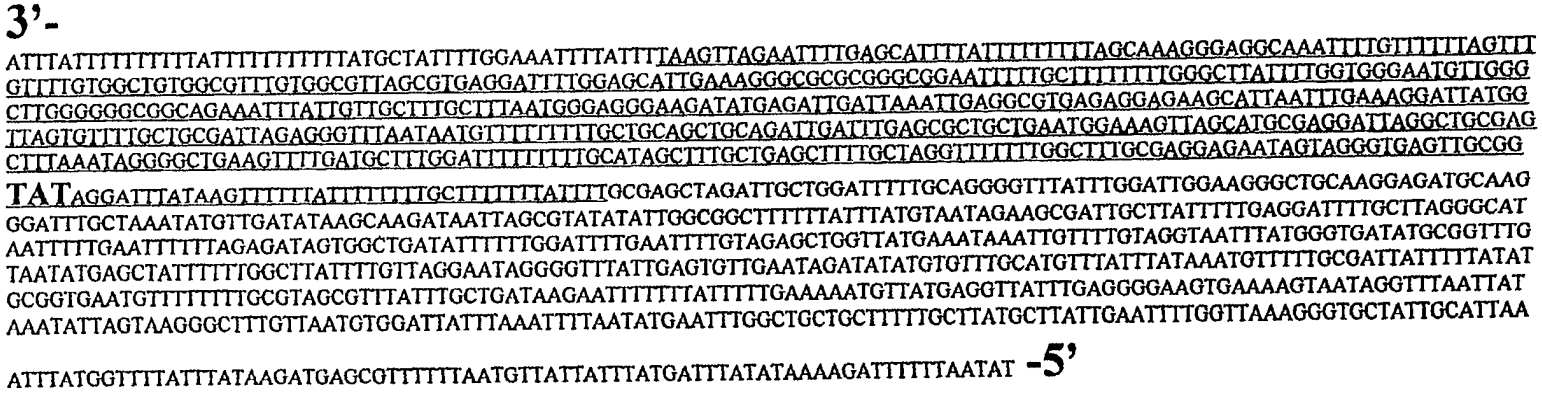

3'-
ATTTATTTTTTTTTTATTTTTTTTTTTATGCTATTTTGGAAATTTTATTTTAAGTTAGAATTTTGAGCATTTTATTTTTTTTTAGCAAAGGGAGGCAAATTTTGTTTTTTAGTTT
GTTTTGTGGCTGTGGCGTTTGTGGCGTTAGCGTGAGGATTTTGGAGCATTGAAAGGGCGCGCGGGCGGAATTTTTGCTTTTTTTTGGGCTTATTTTGGTGGGAATGTTGGG
CTTGGGGGGCGGCAGAAATTTATTGTTGCTTTGCTTTAATGGGAGGGAAGATATGAGATTGATTAAATTGAGGCGTGAGAGGAGAAGCATTAATTTGAAAGGATTATGG
TTAGTGTTTTGCTGCGATTAGAGGGTTTAATAATGTTTTTTTTTTGCTGCAGCTGCAGATTGATTTGAGCGCTGCTGAATGGAAAGTTAGCATGCGAGGATTAGGCTGCGAG
CTTTAAATAGGGGCTGAAGTTTTGATGCTTTGGATTTTTTTTTTGCATAGCTTTGCTGAGCTTTTGCTAGGTTTTTTTGGCTTTGCGAGGAGAATAGTAGGGTGAGTTGCGG
TATAGGATTTATAAGTTTTTTATTTTTTTTGCTTTTTTTATTTTGCGAGCTAGATTGCTGGATTTTTGCAGGGGTTTATTTGGATTGGAAGGGCTGCAAGGAGATGCAAG
GGATTTGCTAAATATGTTGATATAAGCAAGATAATTAGCGTATATATTGGCGGCTTTTTTATTTATGTAATAGAAGCGATTGCTTATTTTTGAGGATTTTGCTTAGGGCAT
AATTTTTGAATTTTTTAGAGATAGTGGCTGATATTTTTTGGATTTTGAATTTTGTAGAGCTGGTTATGAAATAAATTGTTTTGTAGGTAATTTATGGGTGATATGCGGTTTG
TAATATGAGCTATTTTTTGGCTTATTTTGTTAGGAATAGGGGTTTATTGAGTGTTGAATAGATATATGTGTTTGCATGTTTATTTATAAATGTTTTTGCGATTATTTTTATAT
GCGGTGAATGTTTTTTTTGCGTAGCGTTTATTTGCTGATAAGAATTTTTTTATTTTTGAAAAATGTTATGAGGTTATTTGAGGGGAAGTGAAAAGTAATAGGTTTAATTAT
AAATATTAGTAAGGGCTTTGTTAATGTGGATTATTTAAATTTTAATATGAATTTGGCTGCTGCTTTTTGCTTATGCTTATTGAATTTTGGTTAAAGGGTGCTATTGCATTAA
ATTTATGGTTTTATTTATAAGATGAGCGTTTTTTAATGTTATTATTTATGATTTATATAAAAGATTTTTTAATAT -5'

Figure 25

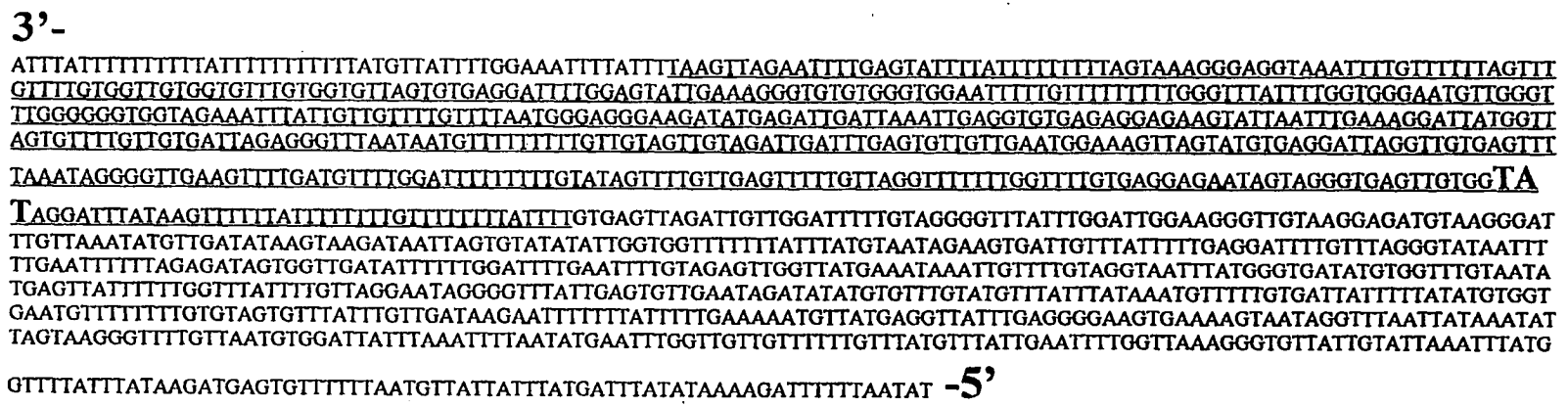

3'-
ATTTATTTTTTTTTTATTTTTTTTTTTTATGTTATTTTGGAAATTTTATTTTAAGTTAGAATTTTGAGTATTTTATTTTTTTTTAGTAAAGGGAGGTAAATTTTGTTTTTTAGTTT
GTTTTGTGGTTGTGGTGTTTGTGGTGTTAGTGTGAGGATTTTGGAGTATTGAAAGGGTGTGTGGGTGGAATTTTTGTTTTTTTTTGGGTTTATTTTGGTGGGAATGTTGGGT
TTGGGGGGTGGTAGAAATTTATTGTTGTTTTGTTTTAATGGGAGGGAAGATATGAGATTGATTAAATTGAGGTGTGAGAGGAGAAGTATTAATTTGAAAGGATTATGGTT
AGTGTTTTGTTGTGATTAGAGGGTTTAATAATGTTTTTTTTTGTTGTAGTTGTAGATTGATTTGAGTGTTGTTGAATGGAAAGTTAGTATGTGAGGATTAGGTTGTGAGTTT
TAAATAGGGGTTGAAGTTTTGATGTTTTGGATTTTTTTTTGTATAGTTTTGTTGAGTTTTTGTTAGGTTTTTTTGGTTTTGTGAGGAGAATAGTAGGGTGAGTTGTGGTA
TAGGATTTATAAGTTTTTTATTTTTTTTGTTTTTTTTATTTTGTGAGTTAGATTGTTGGATTTTTGTAGGGGTTTATTTGGATTGGAAGGGTTGTAAGGAGATGTAAGGGAT
TTGTTAAATATGTTGATATAAGTAAGATAATTAGTGTATATATTGGTGGTTTTTTTATTTATGTAATAGAAGTGATTGTTTATTTTTGAGGATTTTGTTTAGGGTATAATTT
TTGAATTTTTTAGAGATAGTGGTTGATATTTTTTGGATTTTGAATTTTGTAGAGTTGGTTATGAAATAAATTGTTTTGTAGGTAATTTATGGGTGATATGTGGTTTGTAATA
TGAGTTATTTTTTGGTTTATTTTGTTAGGAATAGGGGTTTATTGAGTGTTGAATAGATATATGTGTTTGTATGTTTATTTATAAATGTTTTTGTGATTATTTTTATATGTGGT
GAATGTTTTTTTTGTGTAGTGTTTATTTGTTGATAAGAATTTTTTTATTTTTGAAAAATGTTATGAGGTTATTTGAGGGGAAGTGAAAAGTAATAGGTTTAATTATAAATAT
TAGTAAGGGTTTTGTTAATGTGGATTATTTAAATTTTAATATGAATTTGGTTGTTGTTTTTTGTTTATGTTTATTGAATTTTGGTTAAAGGGTGTTATTGTATTAAATTTATG
GTTTTATTTATAAGATGAGTGTTTTTTTAATGTTATTATTTATGATTTATATAAAAGATTTTTTAATAT -5'

Figure 26

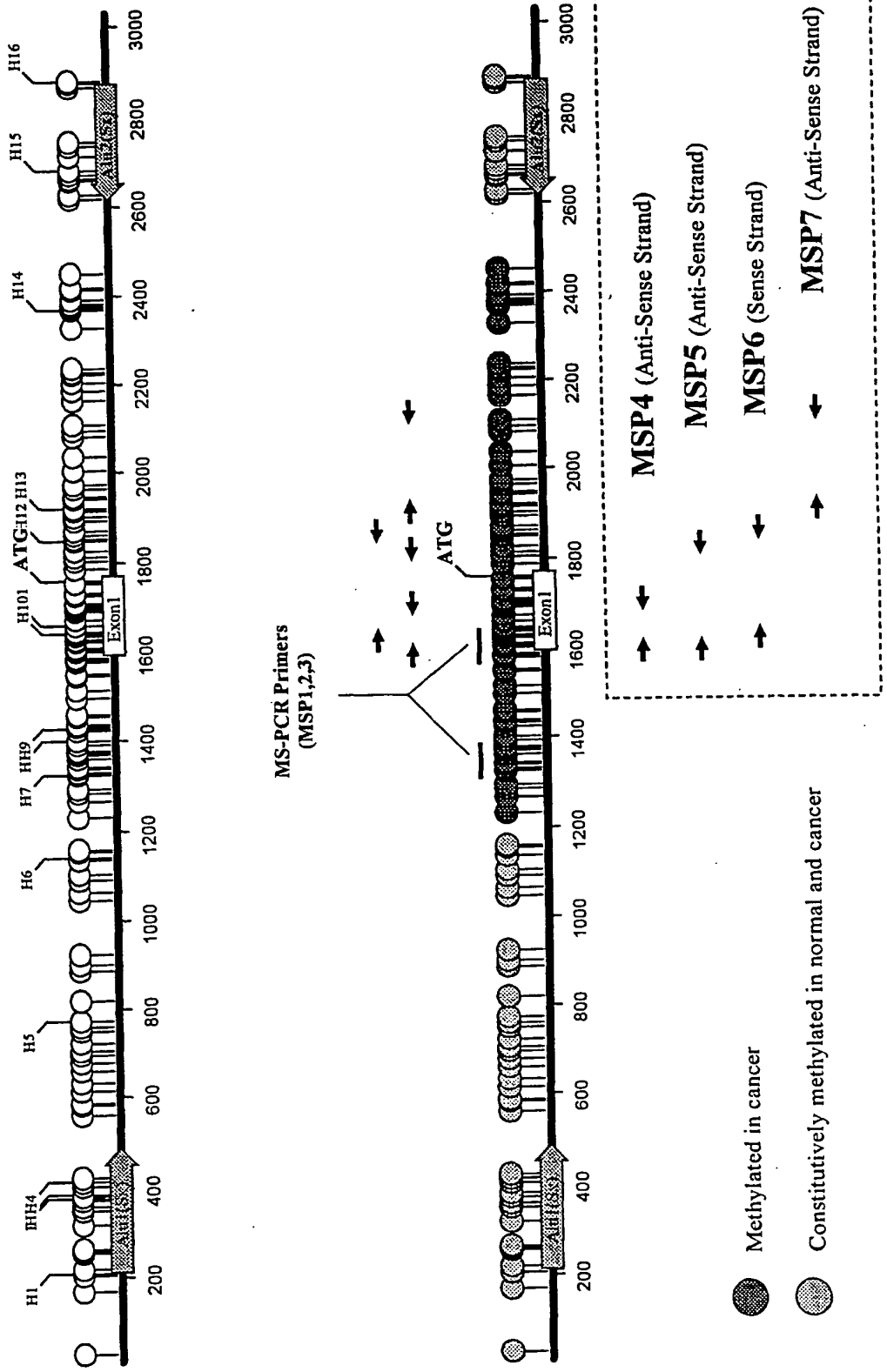

H1
HH4
H5
H6
H7
HH9
H101
ATG
H12
H13
H14
H15
H16
Exon1
Antisense
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
2800
3000
MS-PCR Primers
(MSP1,2,3)
MSP4 (Anti-Sense Strand)
MSP5 (Anti-Sense Strand)
MSP6 (Sense Strand)
MSP7 (Anti-Sense Strand)
Methylated in cancer
Constitutively methylated in normal and cancer

Figure 27

MSP4
(against Genetic-Anti-Sense Strand)

1P-HLTF1581MF(ASS): 5'-ACGTCGACGTCTAACTAAACTCGCGA-3'
13P-HLTF1713MR(ASS): 5'-ATCGTTTTCGAGTCGTTTCGATACGTT-3'
2P-HLTF1575UF(ASS): 5'-AAAACAACATCAACATCTAACTAAACTCACA-3'
14P-HLTF1728UR(ASS): 5'- TTTTGGTTTTTTTGGATTGTTTTTGAGTTGT-3'

MSP5
(against Genetic-Anti-Sense Strand)

1P-HLTF1581MF(ASS): 5'-ACGTCGACGTCTAACTAAACTCGCGA-3'
5P-HLTF1827MR(ASS): 5'- GACGTTTTTAGGTCGTTAGATCGAGC-3'
2P-HLTF1575UF(ASS): 5'-AAAACAACATCAACATCTAACTAAACTCACA-3'
6P-HLTF1829UR(ASS): 5'-GGGGATGTTTTTAGGTTGTTAGATTGAGT-3'

MSP6
(against Genetic-Sense Strand)

3P-HLTF1621MF: 5'-GTCGTGCGTTTTTGATTCGGCGTTC-3'
7P-HLTF1873MR: 5'-GCCCAAAAAACGCAAAAAAACGCCG-3'
4P-HLTF1614UF: 5'-TTTTTTAGTTGTGTGTTTTTGATTTGGTGTTT-3'
8P-HLTF1878UR: 5'-AAATCACCCAAAAAACACAAAAAAACACCA-3'

MSP7
(against Genetic-Anti-Sense Strand)

9P-HLTF1893MF(ASS): 5'-GTTCTATTAATCGCATATATAACCGCCG-3'
ALU(MB)2133FR(ASS): 5'-TTGGGGATAAGGATTGTTTTATTTGGTTT-3'
10P-HLTF1890UF(ASS): 5'-TTCATTCTATTAATCACATATATAACCACCA-3'
ALU(MB)2133FR(ASS): 5'-TTGGGGATAAGGATTGTTTTATTTGGTTT-3'

MSP8
(against Genetic-Anti-Sense Strand)

15P-HLTF2201MF(ASS): 5'-TACGCCACTTACAAAAAAAACGCATCG-3'
11P-HLTF2400MR(ASS): 5'-TTTAAGTTATTCGTATTCGTTTTTCGTCGTC-3'
16P-HLTF2197UF(ASS): 5'-AATATACACCACTTACAAAAAAAACACATCA-3'
12P-HLTF2403UR(ASS): 5'- GGTTTTAAGTTATTTGTATTTGTTTTTTGTTGTT-3'

Figure 28

58381 aggtgagaac tgggcaaaag ttgtgaagca gcaattctgt tatatggaca gtgttctgct
58441 ttttaatcct atttagcttg tttcagaaat tctcactttt gttgactgcc aacatacaaa
58501 gtaagggaaa ctcaagatat taagatggct gtatcagttc ttaaaatctg cagagcctgg
58561 ttcaaaatca gtcactccct tcagaagcag acatggcatc tgttccttgc ttgcttgttg
58621 gttgtgtacc tttcacgaga cctgaatttt agaattgccc agtgctgcca gagtgagtga
58681 gtgtaattct cctttcaggt aaagataggc tatctcaaca ctgctgagtg attcataaac
58741 atatcaacca atagcattaa cccatttttat ttcctgtcct tagtgtctga agatgctcac
58801 cagttttctg tgtacagtaa ggcagcatgc taaaatgctt ttgttcagtt ctgtatattt
58861 gaaaatagca gtgtgttctc tgatggttac ctgcagtggc accctgtaca aaaaataaaa
58921 gacttattgc tgtatcttgg ttgtttaatt aaattaagga atttcaccat acacccttga
58981 acaaatctat tagggaattt ttcacaattt ttggaatttg tcatagtttt aaaaaagtgt
59041 aaagcttgac attgggatat atgctttaaa aactggtatc tatgatttca atctaattgt
59101 ttttctgtga tggtgatgga tctgacagat cagaacaaac cgagatcaaa cttacatagt
59161 gtcatccaca ctgcacctct ctattaagtg ggtatttagc tatttagaat attttaacct
59221 taaagtcatt cctaactgtc agaatcaata ctgggtccag tccctactag tattttatct
59281 atgatttaaa atgatagaat ggaaaagcgt atttattaaa tttatagatc attctaggtt
59341 gtaagagatg gcaagctaaa gaagggtagg atccaagatg acctttttaaa actaaaatga
59401 gtggtcagaa atgattctat ggaagaaaac aggacagtta accatgacag gaagaaactt
59461 tagaggtttg attgggcaaa agctgcaaag cttttgtgac ttagccttca aagtcaccct
59521 ccttattcat attgatgtga attaatatac aagcatatag acagcttaag tcaagaatgg
59581 ttggggccac cttgaaggct gttaccatag gagctcaata agaacctgag gatttaccta
59641 gaattataaa atattaaatg ataaatgact tctgaaacta gctatttgga ctggtgaaga
59701 atgagtcgct attgatcttc aagtacaatg aagcatttac caaagattta tttaagtgcc
59761 tccatgtgtc agatgctgtg gtacaaagaa ggacttctca aaattttagc tagtcagagg
59821 tctttctggc ttccgagtcc ctggttaaga tgaacagaaa cacagtcttc agatataaaa
59881 tgtcttattt ttgtggccat tcagttgcat tcaacgttaa ttttttctat ttactaccgt
59941 tcattctcta ttttttacac agtagcataa caaagctcta aggtggaaaa gctgacatag
60001 ttttaaattt ttttttttt ttttttttt cttgaggcag agtcttgctc tgtcagtcag
60061 tcaccaggct ggagtgcagt ggtgcagtct cggctcactg caacctctgc ctcctggatt
60121 caagagattc tcctacctca gcctcctgag tagctgggac tacaggcacg caccaccacg
60181 cccagctaaa ttttgtattt ttagtagaga caggtttcac catgttgcca ggatggtctc
60241 aatctcttga cctcgtgatc caccggcctt ggcctcccaa agtgctggga ttacaggcgt
60301 gagccactgc acccagcaaa tttttaaatt tcaaataagt agtgaaggct attattaact
60361 tttggaatca gaaagaatga caagcttacc ataagacata gcatataatg ctgtcaagtt
60421 atttggctag aaaatcactg aactaaacat tcttttcctt ctatgatcta tgtcttaagg
60481 tgaagtatta actaactttt ccatgtaaag ctatacaaat attggaaaat cttttctagg
60541 gagtccagaa tactaagggt tacttagtaa aatgtataaa aaggcaacag taattcaaat
60601 tacaagattt atatttgcag aggtgatcca tatatactta tccccttgca gtggctggta
60661 tgacctttgg ttgtaagaca aacttgccca caacagaggt caaatccatg cctttggaga
60721 ttagctccat ggtggatgga gctatggttt atgcataaag taaatgtttg tttaccttaa
60781 ttctccttat acccatattg tcctgctgta taacacattt tgcagatatt ttgaagttaa
60841 tgtgttaaaa acttgaggtt aaacatttga gtttttgtta agagccaaac atcaaatgtg
60901 cccttatatt tttaatgaat ctcatccaaa tgctaatgca taaaccttga caagtagtat
60961 aaataaaaca agaaaaaaat acagcaatgt ctttgccatt ccccaaaaca aagcacacac
61021 tgccgaagat cattagtact cactggtaac aaactacata gggttagttt gtatttccaa
61081 ttctagagct gtaatttttaa ggacaaaatg tacaatgatt gattaagagt gctatctgtg
61141 tatatatagg tattatcaca actccttttt ttcttccaga tgaagaaatt aattgggacc
61201 aatgtttttа gatcaaggca ttttaaataa gcactcttga tttctgaaca agaatttcaa
61261 ccagctaaat tgagcaaaat aaagttagtt aggatatgag gacattattc tgttacagta
61321 atcttcatgt actctcaaaa aaatgtaaca cttgcataga aatgtcacaa ttaatgaagg
61381 attttatttg aagataaagt caaaattatg gcaccgagga aggtaataaa catttgaaat
61441 ttttattgat ttttaaattt aaaatccagt tttaaccaca aaattgtttg aatcacaagt
61501 ggtaatacaa tgtcttcaat atttttctaa agttattttt ctatataata ataagacaac
61561 agcatagcat ataggaagtt ttcattccag tggctttttt atatatttat ccttcttagg
61621 aaggacaaat taaatttttt aaattaaact tttaaaatat aacaacatct aacagaactg
61681 tacaaaacaa agagacattt tttaaacaac ttgccaaact tacttatgag tgtgttttaa
61741 aaacaacttt gtaaatgtct gggcaaagaa gcaagctgtc ctccctttac cttcatagtg
61801 agtttgtaag gctttgtctt tgtaagcaga aagagtagac tgtgttgttt tttgccaaaa

61861 actgtttata cttaatctca ctgaagtatt gctatatgga gaacccatac tctgatcaac
61921 ttgatttttt gtgtgtaatg cttgatctac caggtaactt cccaactgct cctaatgcta
61981 gcgggctaat cccacattat tattccacta tcatccctgc agaaaggtct tggttttgat
62041 gaaaatcagc cctttcctta cctgctactg cctcaaaaag ggaccaggaa gattctagct
62101 ggctaattca ctgttccctt tgagcaagaa aacggcacag ggagaaaagg acttatctgg
62161 tgagagattt ggcatatacc ttcaatgtgt gccctataac acaacattgt ctccgatctc
62221 atctttctat caaatgactt ccaacactct taagtctcag gtattcttaa atctgtatca
62281 tcaaacatga agcttctctt gtttgttaga gtaattaatc tttctttgga ttaaagtttc
62341 cctttgaaat aaaaccacct acctaatctg actgctaaat ttctagcttc tttgttttaa
62401 atatgctcag gagtcaaccc aaattctgca gcaaataagt ttgcttatta acaaaaaagt
62461 aaaaaaaaaa gaaaagaaaa aagatgacta attctacaga tagctgtaag gatgaattac
62521 tcaagttcaa aatcaaattc tgattctaaa cacataacaa ttgtttacat tcaggattaa
62581 gatgtcttta agagttgaaa cgactttgga gatcatccag cccaacttcc atccagatat
62641 cacgcctctc acatatagta gtcttctgaa ttataaaaat ttataaagtt acttccaaaa
62701 aaagctacat aaataaaaat tatctattta tagaaatatc tatttagcag ttccataatt
62761 taaaatattc aaatcaaatt gggtaggact ggtttgcctc tcactcccac agactatatt
62821 tatacctcag acacagcaag ttacatttaa acaatgagtg tagtactact taactaaaat
62881 ggaaaaaata gtactcttaa cataatccct aatttttca tgaacataaa actccaagtc
62941 atttatgtga actatatctc aatgtagctg taggaaaaat aaaaacctgt gctaacctgg
63001 actttggtct catttaagat ttggttctgg aatgcaaata tggtttttga aagcccaata
63061 aaattaattc ttgtatagtc tgtatatatt gtttacaagg actacaaaca ctgcatcaca
63121 aatcggaggc tttggtaaat aactaagtgt ccaacataga aaataactat ttggtcaaaa
63181 gtataaaagg tctgaccttа tttgaaatac gaaaaagctg agtacttgga agatacgtga
63241 aaatactcag catagatatt atgaaaagct gaataacaaa gtaacctttt ttctcaaatt
63301 atttcaggcc acagtatata acggaactta ttgctatttg aagtttcatt aaaaataggt
63361 tcatatatag aagaaattgt gtcagtaata cctcttcact aatataaaat atgccccttt
63421 tagaagacgt gttctctaga tctcatttct aaaactctgt atttttctca tttctaaaac
63481 ttaagtatcc aatcaaactg accttactaa aatcccacaa attataagtc aattaatgtt
63541 ctgatttcat taattttggc ttgtttcatt tcgtcagcat ttggtttttt agttccaaag
63601 gctcctgctg caagttctct ctttttgttt tgtattttca gcatattttc ttcaacagag
63661 tcctttacaa tgaacttтaa aaagaaaaaa aaaagttaag tagtttttaa ggcatagtat
63721 ttaaaatctt tattaaaaaa actaaatgaa agatttgcct aatggcaaaa accgcaatta
63781 cttttgcacc aaactataat aactatttaa tagaactgaa ctttgctttc cttacccaat
63841 atggttttct tatataaaga atatggctgt acgaatagaa gttacagatc acaagataag
63901 caaagcaata aagaaaaaaa aagcaaacta tataggcaac tctggcaaaa tgacttcagc
63961 attgtattaa aaaggccatt ctggccgggc acggtggctc acacctgtaa tcccagcact
64021 ttgggaggcc gaggcgggca gatcatgagg tcaggagatc gagaccatcc tggctaacac
64081 ggtgaaatgc catctctact aaaaaatага aaaagttagc cgggcatggt ggtgggcgcc
64141 tgtagtccca gctactcggg aggctgaggc aggagaatgg tgtgaacccg ggaggcggag
64201 cttgcagtaa gccgacattg cgccactgca ctccagcctg ggtgacagag cgacgtctca
64261 aaaaaaaaaa aaaaaaaaac aaaaaactat tcccagaaat tcaagaatga atctttagta
64321 taattcacaa cactaacaaa gagaaaaata aaacgtgatc atctcattcc tgataaaaac
64381 ttgattttaa aataccccaa ctaggaatga atgtctcaac cttacatatt taatgtgaaa
64441 caataaaagt attctacaac aaaacagaaa caagtcaaga atggtcacta tctgctacta
64501 tttaacattg tttaagaaga cccaattact gcaataacta aagctatata ataaaatttt
64561 actgaaatac atataagatc tacataaatg gagaaaggat actatgttcc tgtatggaaa
64621 cattcgttac ttataaagaa agatagttct cgaaagtaat aataaattca ataaatctca
64681 agatatctta aatgggattt tataactaaa atctaaagtt caattcaaga gcaataataa
64741 taaccaccaa tacaattttc aaaaagggtc atcagaacat tttaccctat caggtatcaa
64801 aacttacttt aggaaaccat agttgttata agagtaatac tgccctaaga aaaggaaaag
64861 agatcaatgg aacagaatag tctagaaaaa gacccaagta atatatataa attagtttat
64921 gatgaaagtg acatttcata ctagtaggga aaaagttgga attttcaaag aaaaaaaatg
64981 atgtgacaga acaaatcgtt cttcatctga taaaataata agcttagact catgtcacaa
65041 atcataaaac aaaaaataaa atcttccaaa agaaccagaa gtaaatacag aatagttgaa
65101 aaaatgtgaa aataaagcct tctgtcctag gtaagacctt aaaacctaga acgtataaaa
65161 aaggctgacg ttttaaatca taaaaatttt aaagcttttt catgaccaaa aaaatgcagt
65221 ttaaagacaa ggaagagtat gggataaacc attttcaaca tacataaaca aaagatttca
65281 agtcagactg tatcagggag tcctttatac tataaaacta tttgtaccat aactatgaga
65341 actggacaga agatataaat aattcacaga aaaaatgtta ataaacaaga tatgtgaaat
65401 atactcaagg ttttaggtaa atgcaaatta gaacagcact tttttggcca cagaagcata
65461 aattaacaat tttcatataa atcctatttt gaaaccaact cgatagttga taaaaattct

Figure 28 Cont'd

65521 aaatccagaa ttctagtttt cttaacttat cccaaagaat tacttgcatc tatgtgaaaa
65581 gatttatata taataatgct cattatggca ctgtttttat tagcaaaaaa cttaatgggt
65641 taaatatatt acgatatatc tatatggttt tgtagagtag tcaaattcag agaaagaaag
65701 tagaatggta gttgccaggg gctgggggga gaggaatggg aagttattat ttaatggata
65761 cacagtttca gtttaagatg atgaaaaagt tctggagatg gatgattgtg atggctgttg
65821 cacaacaaca tgaatgtact taatgcatct aaactgtata cttcaaaatg gtaaattttg
65881 tattttacca ctatttaaaa aaagtgttca tctaaatgta cttatgaggt agtctccaat
65941 gtatagtttt aagggaaaca aatatgttgt agaataaagt ataattccat ttatctttaa
66001 aagaatgttt ttaaggagat gtacacataa ctataaatgt aggtgcagag aaaggcccct
66061 taaacagcat atattaaatt attgatgaag gaggaaaaag gtgacttata ttttgctctg
66121 cataattaca ctcaatcata ataatatatt catgcattac ttgcttaatt gtaaaaatcg
66181 tatcaaccta gtcttaaaat agtttgttta aaaactcact tttgtgatga taacttcttg
66241 cttctgacca agtctatggc atctgtcaaa gcactgatct tcagcagcag gattccaggc
66301 ctaacaagaa catggatgag ttactttact actgccctga tctttcagtg ttttgcattt
66361 gtcctttcat ttcactagta ttcattcatc ttttttctgac tgaaagtttt ctgcctgaag
66421 tcagatacta acatcacaaa tactctgggg aaaaatccac cctcccaatc actatgatga
66481 taggcatcct ccccctacta ctacttctaa tacctaatac aggccccgac atatgtatca
66541 gcacccaaaa taagtatttg ttgattaatc aatttggttt tttttttcat ttttaaaaga
66601 atctatgtgt gcttaacagc agaaaaaggt tatacatcct cttaaatgac ttaatttata
66661 tgagaaaaga cacaattctt gggcctattt agttttctcc ccctttctg gttttctttc
66721 actttcctct gttcatgaaa taaaagtata aaaaaagggt ttctcctggc taacacagtg
66781 aaaccccgtc tctactaaaa atagaaaaaa ttagccggga gaggtggcgg gcgcctgtag
66841 tcccagctac tcaggaggct gaggcaggag aatggcctga accccgggag gcggagcctg
66901 cagtgagccg agatcgcgcc actgcactcc agactgggcg acagcgagac tccgtctcaa
66961 aaaaaaaaaa aaaaaaaaaa aaaaaggggg ttttaagagt cagtaagtcc taaaaattat
67021 ttttttgttg atgctatttt ctgtatcagt atttgacgtt agtttgactt taaagctgag
67081 ctaaaaaaag taccattagc taattaacac agacatgatc atgagaactg ttttaagcta
67141 aaaaataaaa tgatgggtcc ggcacagtgg ctcacgcctg taatcccagc actttgggag
67201 gccgaggcag gcagatcacg agttcaggag atcgagacca tcctggctaa cacggtgaaa
67261 ccccatctct actaaaaata caaaaaatta gccgggcgtg gtggcaggcg cctgtagtcc
67321 cagctactcg ggaggctgag gcaggagaat ggcgtgaacc caggaggcgg agctggcagt
67381 gagccgagat cgcgccactg cactccagcc tgggcgacag agcgagacta catctcaaaa
67441 aaaaacaaaa aacaaaaaac acaacaacag caataaaatg gtgcaacaat gaaaagggaa
67501 gaataaccaa aaagatatag acccctggta agtggttttt aaatagtatc atgaattgtt
67561 tggataagga tcattaaatt ttaaattaaa actatgaggt ttttttttg agatggagtc
67621 tcgctctgtc accaggctgg agtgcagtgg cacgatctca gctcagtgca acctccacct
67681 cctgggttca agcgattctc ctgcctcagc ctcccaagta gctgggatta caggcgcttg
67741 ccaccatgcc ccgctaattt ttgtattttt agtagagacg gggtttcacc atgttggcca
67801 agatggtctc aatctcttga cctcatgatc cgcccacctc ggcctcccac aaagtgctgg
67861 gattacaggc atgagcctcc gcgccaggcc atattacccc aatacttata caatgttcaa
67921 tgaaaaggta aaagaaagct gggcacagtg ccatatgcct gtaatcccag ctactcaaga
67981 ggctgaggga ggattgcttg agctcaggag tttgagacca gcctgggcaa aagaggaaga
68041 ccctgtctct taaaaaaaga aaaaaaaatt taatggtaaa agactttcat gagcaaattt
68101 ttatttttt attgaatcat tatgtttctc ggacctgact tttaagttag aaacaacatt
68161 ggaaagcttg atatgttcaa tgaagaacac tcaacttact ataaaactgc taaaacgttc
68221 tgtaaatttt taaaaactgt taacacgttt agtctctgat caacttttc atactaatta
68281 ttttggtatt aacctctttg gtactaagat tataagcaag aaaaataaaa tttgggattt
68341 gttctaaaaa tcagcttcta gaattatcac tggattaatt aatattcaac aatctatagt
68401 tacaaatatt ttttatagta gccataaaat actgcaaaca aagatttcaa acttcatcca
68461 cagtacaaag gcaagaaaat acaacagacc ggctgggcgc ggtggctcat gcctgtaatc
68521 ccagcacttt gggaggctga ggcaggtgga tcatgaggtc aggagttcaa gaccagcctg
68581 gccaacatgg taaaaccccg tctctactaa aaatgcaaaa attagccacg cgtggtggca
68641 ggtgcctgta atcccagcta ctcaggaggc tgaggcagtg aaccgcttga acccaggagg
68701 tggaggttgc agtgagccaa gattgtgcca ccacactcca gcatgggcga cagagcaaga
68761 ctctgtctca aaaaaaaaaa aaaaagaaaa aagaaaatac aacagaccat accatatcca
68821 tgatattcac ccatgaaggc ctaagtatga tgaccctagt tataatctta ctgagaagga
68881 taccgtcagt tgagggagga aaaaggaagc ctgagtaaac atatatgcat caacaaggta
68941 gcctcaccca agtaactaaa atttatatct cagtagtatt cctacatgta tctcagtaga
69001 attcctacat gttaaaaatc ttctgagttt caggtaacaa gttacaacca caaatcttaa
69061 gtctgaatgc acagaaacag cagagggttg cctttagccc acatagtacc cagtgcaaat
69121 tagaaactga attagtccaa ggtcctctca ggccaaacag ccccttttgta cacgtagaaa

Figure 28 Cont'd

69181 aacacgctcc tatgtgaatg cacgtccagg tcacacagct tggtgagtaa agttcaaagt
69241 gaattttagc ctctgttcta ccttcatgcc tgtactacac aaccacaact gactgcacta
69301 tattaggcat taaaacgtac tcagaacatt ccattaatga taatttgttt ctcaaattgt
69361 gatcctagac cacctgtacc tcaatcacct atgatatatg ctaaaaatgc agcttcctgg
69421 gtctaatcct agaactccta agtctctacg ttgaatgcaa aggaaattgc attttaccta
69481 attcttatgc acattaaagc ttaagactac tagaaccgtg atggtaatca ttttgtctat
69541 cttgtttata gttagattcc cagtgcctat aaaaactgac gcaaagcagg atctcattat
69601 ttgctggctg aatgaaccac taccctaaaa agcattacag ttaatattta ttatagaaga
69661 aaagtagaat agcatagcat aatagtactt aaatatttac atggactatg attctggttt
69721 tgcctattac taactgtatg atactgggca cattatttac ttttctgtgc ttcagttgcc
69781 tcttcagtaa aacagagcat ttttttcttt ttttttcttt cagagacggg ttcttgctct
69841 gtcacccagg ctggagtgca gtgacaggag catggctcac tgaaacttca acctcctggg
69901 cccaagcaat caatctttcc gacttagccc tgcaagtagc tgggactaca agcacaaacc
69961 actacaccag agtaactttt ctggtacttt tttttgtaga gatggggttt tgccatgttg
70021 cctaggctgc tctcaaactc ctgagctcag gcaatccatt caccttggcc tcccaaagtg
70081 ctgggattac aggcatgtgc cacagcacct ggccaaaaca gagcatttaa taatggtatc
70141 tacctcataa ggatgttgtg aggatttcag taaagtgctt aagagaaagc tcttaatcca
70201 catttactaa gtttttgtta tttatttcta aagtttgcct aagaattttc acagttaatt
70261 ataactatga aggtatttac caaatagttt tttgtcttat ttttttgaaa gaagtactaa
70321 gattctgaaa aaatttacag aactacttac tggatccatt aaaaacactc gagaagctgc
70381 agacagattc aaaccaactc cacctgcttt taaggacaga agcattatag ttggagatcc
70441 tgcttcagtg ttttgaaaac actgaattga ttcaactctt ttcttttggg ccatggaacc
70501 atccaaacga gtaaacacaa atccagaggc tctaaagggg ggaagaaaag agacaagtaa
70561 caaacactat tattataaaa ataaaacaaa gttaagtaac tacaaatctc ttcaataaat
70621 tctttcactc atttatttta acaaatgttt attgaacact tattaatatg ctaggcactg
70681 agtcagtcat tttctgataa attcaggact taatctttaa attcttgtgc tgccaggaag
70741 tattctctat tgtgactctg aatatctttc agataaaatg actaaaactg gcaaaacaaa
70801 ttaagttttt ttttttgcta tccaaatcaa tatgattata taaaatcctt ttactgaaag
70861 tgaaaaacac tgtggaactt acttaagtgg tatttctatt aaagacagga atgttgtaaa
70921 ctgagaaaca accaaacttt ttatgttggg attcttcttt cttaagtcag tcaatgcgtg
70981 cattagcgca ttaatctgca aaaaatatta agatgtccat tagattccat tttaaatcag
71041 taaaattagt ctaaaagttt ttatcataaa gaaaaggtcc tgaaatgtta tttccctaaa
71101 atgcaaaaga ctactcattt aaaacaaaaa attaagactg aaagaacact ctaataatct
71161 gaacccattc catttcccct caaattcacc actaaaaata aaaagccagt ggtcaacaac
71221 agaacactaa aatgacagga atgaaagcag aattacgatt ttgatatact gtgtatattt
71281 ccttatataa aggtaaaaca aatactcaaa tatttgcaca tgagtacttt acctttgaac
71341 tggatgtcca ttccatatca gacttttttct cactgtcacg tgctaattct tctggaggac
71401 attctaataa attatcttca tgtatatcat ttctgcataa agggcattta gcatgtggct
71461 atataagaaa gaacgaagta tgagcaacac ttaacagaag aacaaatata aatatgccta
71521 aaaatttttа aatgtatgct aaaatctgta tctctcatag ctaacatttt attaaaccttt
71581 ttaatggatt tcatccattt aagagacatc tgaacatgct gaactccaga cataatgtgg
71641 taagtagaac ggataaagga gggagcaaaa gtatgtgatc tggactagct agaagcaaga
71701 gcaagaaacc aaaaattata acattgtaag caagtttaaa aaaaaaatct atgaataaca
71761 atttattgaa cctaaataca atacagctac actatttagt actattcttt aaaaaaaaat
71821 ccatcttaat ggcttataaa gctattaagc ttataaagct tataatagct tataaagcta
71881 ctaggatctg ttaaatccta tcaagttcat tccaaagcaa catctgttgt acccgtttga
71941 acctgttaat aaaataagat tatatcttaa ataactatat ataaaacacc attacaatta
72001 aataattctg aagctaaaat tatttaaaag tcaatacagt ggggttactt ttggggaaga
72061 gggagagggt aaatggttgg gaggggccac aaggggacct aggcagcaga tacaggaata
72121 tgtttacctc ataattgact gagctattta attatgaatt tatacatttt ttctacatat
72181 gttatatatc aatttaaaat gtttaaaaaa attctacccc agctgaatga ctgcttaaca
72241 gactagtcaa cttcaaaaaa attctaccac acaaaagaat ctagacctac cttatccaac
72301 acaatagcca ataatcacat ggggcaatta cgtccttgaa atgtggctaa gccaaattga
72361 ggtgtgctgt aagtatacac caccaccaaa ttctgaagat ttagtttttt gtaaaaagca
72421 cattaatact ttttatattt tagatatatt gggttaaata caaatgtatt aaaattagct
72481 taatgtttct ttttacttaa gagtatgact actagaaaac tgaaaattat ctttgtgtct
72541 cacattatca atccatttat tgaacagcaa tgatataaac agcggaatta agatattgac
72601 tgtgaacatc tttttatgtt gcaaatgttg cattagaaga aatatatcac ctattaaaat
72661 actggaataa ttaactatca aaggtatatt attaaatcca agtgcaaaac tccatatact
72721 atcttttatt atattaaaga caggtacact actctatcaa acactgaacc tgtactgagc
72781 aaaaaactaa cctgctcatt ctgaatgact tggcaaatac agggtttaca aaatacatgt

Figure 28 Cont'd

72841 gcacaatgtg ttatcacagg aactgttaaa gaatccaggc aaattgcaca ttcctcatct
72901 gaacctgagc tcagaattaa cttcatcttc cttattaact tctttctcag ttcttcaggt
72961 gtatcatttc ctagagaaaa ggctgaaaaa ttaatttcag agcaaggttt gtaatatgac
73021 aagttaatct atcttataag gaaaagtttc gcttgccagt agggaaagtc tctcatcatt
73081 ttctgccata ggcaaaaagc ataatttaaa tttgaaattt aagaaacatt cagaagccaa
73141 aatagtagcc aactcacata atatacctca tagagaagca tcattataac atgacaataa
73201 cataccacca cacaataaaa ggcagatcca atatgaatct ttcatcaaag ggtatctatc
73261 accggaattc ctggaaagac ctcattataa ttccatttca ttggtaaaaa aattattgga
73321 tgtattccta gctgagtctc acactaaaag caatctccat ttgacagaac aaaataaact
73381 aaatgtatgc cacataaaaa aaatttatca ctcttattta agtatctctg acaaatacaa
73441 tgatatgaag caatcaaatt tacctacctg aggggccatt ggaagacact gcatttgtaa
73501 gaaggtaagt atggcaacaa atttgccgca gtctaagcaa aagacccagg acatctgcat
73561 aatgtgccag gacagtccct tcattaaaat acctagagat taaaatgtca aatattatta
73621 agtccgctaa acaataataa cttcctattc taaaacagca gagtaaaatg gcatttttct
73681 tttcttctct agaaattact ccaaaacaag aatgaaaagc agaaaagcaa ctctaataaa
73741 accaagagag atctataaca ataaggcaca atatatgaga aatggctgcc aattgcagtc
73801 taagtcagat aggaagatgc tgagagggat gactgcagat ctcaaacaga gctgccaaaa
73861 cacaattctt taaaatggat ggcatgccct gaggggaaaa tcaatacctc tcatgtgcag
73921 aaacaggaac agggtgcatg cttggtgact tgagcttcca tggagaaatt tatatcaaaa
73981 aaagagggaa ggtaaactga atacagaaac atacaggtgc actgtatttt caaaaaaaga
74041 ctatgggcac agcactctac actccaagca cccctaaagt ttctgatctc tgtttatgtg
74101 tactaataaa acctaaagta actcaacatg taaccctgag tcttaagtaa attgggttat
74161 ttcctgctgg tcggggacat gattgtgacc tttccctttt gtatatctgg actagcaaat
74221 agctagtcca agaagaactc aagatattca aatataagta ataatcagca gagagcaaag
74281 gtaggggaca catggcagtg aaagaaaact cagcagaaaa atgttgccac agagcaaata
74341 aaaattatga ccaaggatac tgtcacaatt taacaccact caataaaaca attcctccaa
74401 aaaaaaaaaa tcttaaagca aagacaaaga gctttgaaaa aatatggcaa gagaacagga
74461 agagatgaaa gatgagctgg tatggctcaa gaaagtgaaa gaaaacagtg aatcatcaga
74521 gaaatgaaag tcacactgaa accaatataa gccaatataa aaaaagacta gtcactgata
74581 aaaacataga aaaggatgta ggaaacagga ctgagaaaat caaaattaaa caaaaaaatg
74641 taaaagatca gaaataaaat gatacctatg gaacacagat aaaagacatc taacatacac
74701 agagagaaat aaaataatgg aacagaaaac aatatttaaa gaaagctttc caggaaagag
74761 aaggggggaa aagcctggaa tctaaagatt taaataatat cacatgtccc aggggaaagt
74821 caacaaagaa gttgctgaac ttcagaggga aaaaaaaaaa aaaaaaaaaa accaggcttc
74881 cccacagaaa cattctatac aaaatacagt ggtacaatga caacatagat tcttaaggaa
74941 atgggtgatc caaaaattat ataaccaggc catctgatat ttaagtattt taaggcaaca
75001 gaagaaatat agcctaacgg ctaagagtac aaatttttag agctagctgg tcatgaattc
75061 aaaacctggc tgtgccaatt taattatgta accttggaca ggttgcttaa cctatgtctt
75121 gagcttcctc agttctaaca gtggcattgt aaaagtacct acctcataca actactgtta
75181 tttaatttaa tacatgtaaa gtatttaaaa tagtgcctgg actatattaa gtgctataca
75241 tgaatctgtt gctagtatta gtactggtag tattatggca acaaacattt ctgaatattc
75301 aagaatgcaa ggaataggca aaataataca aagattaaac ttacaagctc taaaccaaag
75361 tgccagagtt aatctgtcct gtcactcact agctgtgtga caacagtgta agtcacctaa
75421 actgctctgt gcctcagttt cctcacttat aagatgtgga taattatact gacctaatag
75481 gattgctgtg aatacattta aagcactaag atcaatgtat aacacctagt gagtgctaaa
75541 taaatgttag taaatatata gcctattaaa aagaatactt taaaaataac ttatttttctg
75601 agaacagttt caataaagtt aatctttttc aacttttca aatagaataa cctttactaa
75661 cactaattat tccctggcat ggtcatatct tactgatact cagcatttca atttaagaga
75721 tctcttttga gtgcctacca tgtgctaaat attagaagct tccttcctcc tatttcacca
75781 gccattgcta tcttcacagt tagatctctt tcactttact cactcccttg gtgatctcat
75841 ggctttaatt accatccaca catagctgac tttgaactcg actctcctat ccaactaatt
75901 atttgatatc tctacttgga aaattcacgg gcatcctatg atcaatgtgt ccaaaactga
75961 gcttctgatc ttagcttaaa aaaatctctg ctcttcccat agacctgtct tggtaaacag
76021 caattccagc cttccactta tgcaggtaaa aaggcttgaa gttatctttg attcctcttt
76081 cacatgccat aacatatcaa ccagcaagtt ctgttggttc tacgttcaca atttaacaga
76141 atctaatcat ttctcactcc ctttatcact actatggtaa aaaccaccat catctcttgc
76201 tttgactgcc actggtctcc ctacttctga ccttattcct attcaaataa cctaaatctg
76261 agcacagcag ccagagtgat cttttaaaaa tataaatcag attacgtccc tctcctggct
76321 caaaaccctt taaggtttcc tgtcataaaa agtaaaagcc aaagtccttg acatatgcgt
76381 tacatacccc cattccaata aatctctgat atcatgacta ttttctctcg ttcatactat
76441 ccagccacac ttgcttctct gttattcttc aagctcacca ggcatgttcc caactgggag

Figure 28 Cont'd

```
76501 catctacact tgctcttcct tctggctgaa acacccctta atcccaatct cctttagatc
76561 tttacctata tatcatcttc tcagtgaggt cttccttaat accctactaa aacagctttg
76621 gcattcccca tcccctttcg ttgctttatt ccttaaaact taatcatgtt tttaaattat
76681 ctgtttcttc ccattaaaat ataaattcct cgagtgcagg cattttgtc tgttttattc
76741 tgctgtatcc caatgggtat aagtgcctgg cacacagtgg gcacttaata aatttttgtt
76801 gaaagaatat gaaaagaaag aagattttga actcagaaat ctgctgctta gaacaatctt
76861 tcccactctc tctaattcct atcaaacctg ctttgtacct acctcactga agtttttatg
76921 ttttgaagat tcccctttca cagtcctatt ctgctttcac ttgtctctta ttcattctgt
76981 tccctgtggg ctcacattaa acctgatctt gcttggacca gaatttcata ccttctacca
77041 atccttaatc ttcttgttaa acctatttca tcagttccta tttccagatt tccatacgat
77101 tttggaaata gaagcaagca aaatcttctg gctagcaatt gagaacttgt catttaaagt
77161 aagctttaaa ttatttccta acttaaaaaa ttacttacct ggtgaattat actattattt
77221 cattcacttt attgatttgt aaaaaacact gtaattcata gaagcaaatt atctttcagt
77281 atacttcaaa ctattgctaa cacatctaaa taatgtcatt cagtgaatgg gaaacaaagt
77341 aaacaaaaac taatttttaa catggttttc tttctccata ccttccaata gtggctctgc
77401 cttcattttt cacagactga taaatctttc tctcttcatc tgaaagtgta atgtgctgaa
77461 taaatacttt acgttctggt aactccaaaa caggttttcc tttaattttg cttgtctttg
77521 ttcttctaag tgtaatattt ttaattaggg actgtaaacg cctaatcaga ataaaacaaa
77581 taattatgta acttttaatt tctacataaa ttgatccaag aagaacgaaa cagaacattt
77641 taatttgtta ctgttttata gcttctgtaa aagtgttttt aaaatccctt atttgggatt
77701 ttaaccaatc tgataatcac agaagtattc tacaaaaata aaattattaa aacaaatgta
77761 tggaaatttc ccctcaatac ctaacatagt ttaatctttt tatatcctgt ccttattttc
77821 gtaagttctg ttccccattt ttgttccaat cactaactct tttacttctc aactgtcctg
77881 tcaactccat caccgaatag tcaacactca aattctaagt atctattaag tgccaggcac
77941 tgaattaaat gttgtaagaa tacaaatatg ggtaagatac tacattctgt aatattaaaa
78001 tcaaaagtat ataaataagt aattgtataa agtaatgagt gcccccaaaa taaaaagata
78061 atatggtaat tcagaacaca gaaaaatgct ttacaggttc cacaatcttt acatttcttt
78121 ctcagtactg gaaaaaaaga gatgtatcaa ttcttgtttt ttcccaacat tgaaatataa
78181 atacaatttc ctggagttta aatagttggt tttgaggaat caagtggata gtactgccaa
78241 ctaggttagt ggctcttgaa gctttctcac tataactcac agtaagaaac acattctatc
78301 cagcaaaaat atacaaacac atacaaaaga aacaaatatt tcacaaaata atacttacta
78361 tgtatcatgc actccgatat ttactaatct cttctctatt tcattgtctt tttaaatgct
78421 ggctaccacc tttataatga ttttatatcc cactaatagc tctcaaatct tcagtctata
78481 aaccggtgct ttagccaaaa agttgaacca cattacatct tgcaaaatgg ctatcactga
78541 gttttaattt atgttctcaa cttttatcta tctccaccag aaacaaccac ataaaacttt
78601 cctatacaac ttttgctaag gtcatggtac ctcacatcat tcactatata aagctatgta
78661 aggccaggtg tggtagctca tgcctgtaat cccagcactt tgggaggtcg aggcaagagg
78721 ataacctgag gtcagaagtt tgagactagc ctggccaaca tggcgaaacc ccatctctac
78781 taaaaatata aaaattagcg gggcatgctg gcgcagacct gtaatctcag ctacttggga
78841 ggctgagaca ggagaatcgc ttgaacccag gaggcggagg ttgcagtgag ccaagatcgc
78901 gccactgccc tccagcatgg acgacagagc aagactccat ctcaaataaa taaatacata
78961 aataaataaa taaagctata taagatatga aaacttaatg catattagat cctcaaagga
79021 tatacgatgt ttcctgctaa atctgttaat gctggttttg gggctttggg gcttagcaac
79081 acacaaaaaa cagttatgtc tttcctttca gtttagccac aacaggggaa gggtagttat
79141 cgacaattag aaatctttgt ctttccaaac aagcatcact aggttctgaa agacaaaggt
79201 actgagccat aaacacttt tagaaggaaa agttaagagt caagttaaaa tagtaaagaa
79261 agtgccaact ggttcaagct actaacagtt atttgtaact aatattaatc actgtctgaa
79321 agtacttacc taagtcctcc ttcatctccc attgtgacag gacgctgtat tgttctatgc
79381 caccattctc tatcaataaa tggtttaagt tttaaaaagg aaagaagaga ccacaagtcc
79441 tttaaagaat tctggattgg agtacctaga aataacagga aactgttata actctttaac
79501 cagagtatcc agtaagagta tctatttggc ccaatcagag aagacaacaa ttgaaaatgg
79561 gcactctgta ccaggctcat cattatcaat taaattaaac tatcaagtgt tgactataat
79621 catttattaa aaatatacac atctcaagat acagttttgc tctaattttt agctctcaca
79681 tgacacacta gccaagaaaa agactggaaa attcaacttt ctccttatca agtacatctg
79741 cttatttagt ttttaaagca gcagcactag cagaaggaaa aagcaactag tttaagttta
79801 cattcttttc cttaccttct agtaaaggta ctgtgaatgg caaacaggtt tgaggatggg
79861 aaacagaaat caagattatg actggatgac tagagtgtct actccatttc cactagagtg
79921 tctactccat ttccaaggtt caagaataag tctagtgact gttccataaa ttatcaataa
79981 gtcaaatgtt gttatatcac gggcaattat gaagaacaac tcaaaacatt tgattaaatt
80041 atttatatga ttaattttgg ggcagaattt acacccactt taagtttaca aagttacttt
80101 actaacttac ataatgatat attttagaga tttaaggttt tagtaagttt aatgctatga
```

Figure 28 Cont'd

```
80161 ttacctgtca aaacccatct tctttctgat tctaagtcaa gtacagcttt tgtctgctga
80221 gcatttggat ttcgtatggc atgtccttca tccaggatca ctcttagcca ccttatgcta
80281 tgtaatggac tatctccttt agtctgaaat aaatgtttta tatgaattaa aaaacacagg
80341 aaagtaaaat agtacttaat atgatttgct aactagttgt tatcattatt taatatcaaa
80401 tttactatgt gctaggtact gctttaagca ctacattaat tattgaatac tcacaagttt
80461 ataaaataag tacttattag cacacctaat tttacagata agaaaactga agcacaaact
80521 gctaagtaag tatgttaccc aaagcaccac aggtggcaga actagcattc aaacccaagc
80581 agcagtctag ctctctgttg agtcactggt caatacactg aactgttcta cctcggatta
80641 aaacatagca atgatataga aaatggcagt tgtaagaatt acttgaattc aacattactt
80701 tcttttatat aaaaggtaat atttttacta gaagactgaa attttgatag caaggtattt
80761 ttcaaaaagt tttaactttt gaacaattgg aattataaag gtaaaaaaca attagactaa
80821 ctggaaaact tatgaactgt aactacaaaa aaagcatacc ttcaaactga gatattaatc
80881 aaatcattgt gatagaaaac aattttccag aatattaaaa aagaaaataa aattcattta
80941 caaattgctt catatactta caatccaaat ttaaaatctt aacttaaaat gttctaaatg
81001 aaaaatattc ctgtctgggt gtggtggcac acgcctgtaa tcccagcact ttgggaggcc
81061 aaggtgggtg gatcacctga ggtcaggagt tcaagaccag cctgaccaag atggtgaaac
81121 ctcatctctg ctgaataaca aaaattagct gagtgtggtg atgcatgcct gtaatcccag
81181 ctacttggga ggctgaggca ggagaatcat ctgaaaccag gaggcagagg ttgcagtgag
81241 ctgagatcac accattatat tccagcctgg gcaacaagag caaaactcca tcgagaggga
81301 gggggagggg aggagtggag aaaagaaaaa tactcctaaa ttagcatcta aataaaatgc
81361 aacccttca cgcaaaaagt caaaaaaaaa aaccatttat ttttctatat atcgttaagt
81421 atctaaagag aaaaaaaaag tataaagcct gaattttgga acactctaga aacctgtggc
81481 atatttctta cattcaatct ttaaaagatc tgttaagtat caacaatact tactccatag
81541 tcatgagtta aaatattata cgtagtcaaa acaatatcct gttttgaaag taaggccggt
81601 tctctaatac gatcaggacc ataataaaca taaaaattca agtgtacatc tgattttata
81661 tgttgtccaa actggtccta aagaaaaatt aggaatattt ttaacaatga gcagatttgt
81721 gtcagactta atagatgtat aaaaaagtaa ctgccctggc agtagtgtga caaatccaac
81781 tactaaaatt gtatgaaatc aagtttcctt cagtctagac caaaataata acaataagaa
81841 aaatgacaca tacacagtaa taataaatta tttaaaattt acttggccct accaatcagc
81901 aactgcttca tttttaggag gaaaaaccaa agttattgaa aacaataatg agttttataa
81961 aaatatttta attagaaggc cctttaatgc tgcattagtt agcagactta ataatataaa
82021 aattaaacag taaaagaaca acaaccaata tttccaagga tgtgggcaga gattcaagca
82081 ctgtagccac tttctttcac tcatttatta ttggcctact atttgccagg tatcaaccac
82141 aatagtgatg taactaaaaa tgaagaacaa ataatcatcc ctgactttct tgaggacatg
82201 atcagtttta aaaacttatt attatttttt taatcatgaa aaaaaccaga cctagaagat
82261 tttcagcaga aaaaaatagt ctgattaaac tgaaaagaga agataccagg gagacacagc
82321 tgggaagcta tgaggagaga gatgatgaag tcctaaacca gcaaggtaga agtaggaaca
82381 gggtggaaag gacaggtgtg aaaaacagaa actaagacaa gactgatagg acttggggac
82441 taatgaaata tgaaatatga ataataataa taagactaag atgctctgat acttgaatga
82501 ctacctagaa ttgtgtgata aaaatataag gaagtttaaa aaaaaaaaaa aagaagggaa
82561 aagtgggcat ccaagtaaga aagagatgaa tattttagac ctttatcatt tttggtcact
82621 atagaatacc taactgaaga cgtatactag gaaaatgtaa ataattacct atagatagaa
82681 gtgagaggtg tagatatgaa ttttaaagtt aacaacagac agtatttaca gctgtaagaa
82741 tgaatgaaat ttcctaaaga gagcatgtac agtgagcaaa cgttagggac agagtctaca
82801 gaataccggt atgtaaggag aagacttgtt tttattggga gaaaacacat gatgcctggg
82861 gtataattaa aaatattcca ggccaggtgc agtagctcat gcctgtaatc ccggcacttt
82921 gggaggccaa ggcgggtgga tcacttgagg tcaggagttc gagaccagcc tagccaacat
82981 ggcgaaaccc tgcctctact aaaaatacaa aaattagctg ggcatggtgg tgatggtgat
83041 gcgcacctat aattccagct actcgggagg ctaaggcatg agaatcactt gaacccagga
83101 gacagaggtt gcagtgagct gagactgtgc cactgcactc cagcctgggt gatggaatga
83161 gattctttct ttaaaaaagg ccgggtgcag tggctcacac ctgttaattc caacactggg
83221 aggccgaggt gggcggaacg aggtcaggag ttcgagacca gcctgatcaa catggtgaaa
83281 ccccatctct actaaaaaaa caaaaattag ccgggcatgg tggcatgtgc ctgtaattcc
83341 acctactcaa gaggctgagg cagcaggatc acttgaacct gggaggtgga ggttgcagtg
83401 agctgagatc gcgccattgc actccagcct gggcgacaga gcaagactgt aaaaaaaaaa
83461 aaaaaaaaaa aaatcaatcc agcaaagaaa aaaaggaaag cacatgtgtc agacttttgg
83521 taatttttga atcttgatga gtagatggag atctcattaa aatttttaa aatactcttt
83581 ttaaaaaatt atatgttgta tggaaataca gaaaagacag atttaatagg aaccagaaca
83641 aagacagaga aaaagttagg acaacaaata taaagggtaa aggcctggga tccaaagaat
83701 ataaaagttt caagaaaaaa aaccaaggtg gacttttaat ttgtcatcta tcttcattcc
83761 agttctttac aagaaggttt agagatagct tacaaaataa gacataacaa ataataaatg
```

Figure 28 Cont'd

```
83821 aataacatat tcaaggcaga ggaaaagggg agctcattgt acatggcaat attgaagtct
83881 gaagatctct gcaaagggag tttctgtagt gagtggaagc aaaaaccaga ctaccaagga
83941 ctgaagaaag tattaaacat gaaataaagg caaaaatgca gagaggataa aacatgaaaa
84001 agtaaagctg aaaatgcaaa gaatttctta tagaaatttg gtaaagaaga aaaggaaatt
84061 agaatccagc atgacaaaga atccaatatg agaaaactga atcaatgggt ggccttgaaa
84121 tgggacagtg cattatttgt aaagggatta cacttggaac agaagaaaat cagcttttct
84181 ggagaataaa gggagagaac tatagttact ccatctctag acagtgttag gctctactgt
84241 atatgtatgt atgtatcatc ctctcaatag aaaacgaaat aggtgttatt tcatgagatt
84301 acttgtctaa atatccagaa caaatatgtg gtagaagtag gattggtacc caggttctat
84361 tcctttacct ggattatttg aagacaaaaa agggaatata agtcagttct agcctgctct
84421 gtagaatgct agatccttta cttgagagcc aggaagacgg tatcttgttc agattgtgaa
84481 attttaattc ttgatcattg aaatcgtttg gatgttcatc tcctccaaat ctcatgttga
84541 aatgtaatct ccagagttgg atgaggggcc tggaggtgga aggtgtatgg ctcatagcgg
84601 tggatccctc ctgagtggct tagtgctatc cccttggtga tgagtgagtt ctcactctaa
84661 gtccacacga gatctggttg tttctaagtt tgtagcacct ctccactctc tctcttactc
84721 ccactctggc aatgtgatgc tgggtcccca tcaccttctg ccaagagtgt aagcttcctg
84781 aggctgtgta agcttcctgt ggcctcacca gaagcagatg ccagcacgat acttcctgta
84841 aagcctgcag tatcatgagc caattaaaac tcttaaatta tccagcctca gttatttctt
84901 tatagcaatg caagaatgga ctaacacaac catgaatccc taaatgccaa accacctctc
84961 attcctttaa attcacagag tgtgcatgat ctcacttggc aagattttgc cacactttcc
85021 tttctgaact ttacattact tgaactttcc taaatataaa attataatta aatactgatt
85081 acccatttta aaaatacatc tctatacccc aagccctcaa ataatcagtt agatagttta
85141 tacctccatt aatactctaa acctcatccc aaaactgcca ttcaagtacc catgtactat
85201 ccttcccata acaaacaaca aattaaagaa agaacagttt cgttggaata aaaaaatagc
85261 aacagaccat taacattcat tctttctggtc ctaagaggaa aaatgataca cttatcagga
85321 caaggtgtaa tttaaaaaat aaataaataa accaaaaact tatataaaaa aagataattc
85381 atgtcacgga aaaaaagact ataaagccaa gaagaggcca ctaaatttga caattaaact
85441 atcacaagta gtcttttcca atgcaattta gagaactaat ggtacaagcc agattacaat
85501 gaattaagga attagtaagt aaaagcacaa agtagaaatt aacctccaag agtatgacag
85561 tgaaggaaag cagaagagta gcaactttta aagatgcata tcaagagaat gtattttaaa
85621 atgaaaaaac ctgagtatgt atcaggaaga gccagtggga agagactaaa taaacaaggg
85681 aaaaaagaac agatgagcaa tgtgggaaaa gagagaatcc aatgtagaga ctaaatactg
85741 gtcttaaaaa gagtaggatc tcttcctcct ctgacataag aagaatggcg atgagagata
85801 aacagataac ttggaagtgg caagaacggt gagtggtcct tatattcaca gggattccaa
85861 gagagtcatt atttcatcat aacaagaagt ttgaagctca aactaggaag aaaacatatc
85921 caggctgcaa ataaacaatt aatcaagagg ttccattcca atgtcaatgt ttcagaacag
85981 aaaagcaatc tggaaaaaaa aaatttctga tgggtagtgt taaaaggaca ggctctagat
86041 tcaaaatgtt ggagactata tacaagatcc atcatttatt agctggatga ccttaggcaa
86101 cttaacctct ctgtgtttca atgtgggcac ctggaaaatc agggtaacag ttatgagaat
86161 taaatgagaa aaatccttaa aagaggttaa ggacaaacca aataaacagg agttcactaa
86221 ctgttgttat tgttataaaa ataagttatt gttggtatag ctcttagaaa agacagttcc
86281 aacagtcaat ctgtaaatgt agaggagtca aactatacat ctttcccact ttgggggaca
86341 aaaatagcaa aggcagaaat acaaaccaaa tcttatgtta gcaattgatt atgttccttt
86401 accttcttct ctactacaca tcttcaaaga aaaatttctt caattcagca aacataatcc
86461 aattcacctg atgattattt gaacttagat acgaactgat gaccaagtat atttacccca
86521 atgaatgact ctttaacaga atactttcta ccttgtagat aaaatataaa ttatgttaca
86581 tttttaaagg gcttaaagac ttacaatcca gttgcttaac acagaaagcg gacagatgat
86641 cagtgttgtt cttggtctct cctcaacatc agttttcttt gacccctcca ctgcacaagc
86701 tcctaaaaaa atgaaccact gatagaacct ttagctgttt ttctgaaata tatatgtcaa
86761 taaggagatg gcaataatgt gcctccatga aaagataaat taggaacaat taaaatgaag
86821 cagatggact tttaacaggg aataaatttt acctcatcac ttaaaaaaac taaaaatcca
86881 catagtgatt atgagaagaa ccaaaggact atgacagaga gtaggtcaca aggaggaagt
86941 ctctaatatt ttctgatagt catagtcatt cattaagacg aacatttgtt ttatgcagtt
87001 ttctgaattt gtgttatatt taatttggta gccagccaag atggccctca atgatctcta
87061 cttcctgata ttcacaatct tgtgtaatct cctcccatac tccaccagga ctggtctgtg
87121 tgaccaacag aatatagcag agaagatagt atgtcacttc tgagatagat tacaaaagaa
87181 gtttccgtat tgggtgtgca agctcactct catactcatt ctctctctcc aatctttcac
87241 tctggtggag ccagctgcca cgatgtgagg ccactcaggc aacttatgga aggatgcaca
87301 tagtgaggat ctgaagacta cagttagcca gcaagtaatg aggtctgctg acaactatgt
87361 gagtgaactt ggaaatggat cctccctaag tccaatgttc agatgacggc ggcccctgaa
87421 gaagagtttg actgcaacct cgtaagagac cataagccat taaaccatcc agctaagcca
```

Figure 28 Cont'd

87481 ctccgacttc cagccctgtc aaagtgtgtg agataataaa tgtctgattt taaatggcta
87541 agttttagga taatttgtta cacagtaata ggtaactaat atattacaac aaaaattctt
87601 aatacttttc tgtggagcat gtttaaaaat agtgataata aagtgttttt acaaattctt
87661 gaaataaatt ctagagatgg tatcccttac caatggagta tcttccagta aacttaaaat
87721 gcaaatttct gagtttttaa aaaattgtat tttcacacta atttcaaaat taaactggaa
87781 ggtgagcaaa gtatcatact gtatctacgt atcattcaca acttgctttt ttctctcagt
87841 gctatctaca gtagtccctc cttatccact gggggtatgt tccaagaccc tcaatagata
87901 cttaaaacca taaatagtac caaaccctat atataatgtt ttttccctat atatacctac
87961 aataaagttt aatttataaa gtaggtgcag taagagattc gtaacaacaa agaataaaac
88021 agaacaatta taacaatata cagttgaccc ttaaacaata tgtttgaact atgtgatgca
88081 cttacatgca aatttttgtt aaccaaacgt ggattaaaaa tacagaatgg gccgggcgcg
88141 gtggctcacg cctgtaatcc cagcactttg ggaggccgag gcgggcggat cacgaggtca
88201 ggagctcgag accatcctgg ctaacacggt gaaaccccgt ctctactaaa aatacaaaaa
88261 attagccggg ggaggtggcg ggcgcctgta gtcccagcta ctcgggaggc tgaggcagga
88321 gaatggcgtg aaccccaggg ggcggagcct gcagtgagcc gagattgcgc cactgcactc
88381 cagcctgggc gacagcgaga ctccgtctca aaaaaaaaaa aaaaaaaaaa aaaaaaaata
88441 cagaatgggc caggcatggt ggctcatgcc tgtaatccca gcactttgga agactgaggc
88501 gggtggatca tgaggtcagg agatcaagac catcttggct aacatggtga aacccgtctc
88561 tattaaaaat acaaaaaatt agccgggcat ggtggcaggc gcctgtagtc tcagctactt
88621 gggaggctga gacaggagaa ttgcttgaac ccatgaggtg gaggttgcag tgcgccgtga
88681 tcgcaccact gcactctagc ctgggcgaca gagtgagact ccgtctcaaa acaaacaaac
88741 aaacaaacaa aaaaacagta tctgctgatt gcaaaactca tgtatatgga aggccaactt
88801 ttcatatata ctcgtcctgc agggcagact gtgggacttg agtatgcgca cattttggta
88861 tacacagggg gccctggaac caaatccctg acaaatactg agggatgact gtactgtaat
88921 aaaacttatg tgaatgtggt cattctctgt ctctcaaaat accttattgt atataatatt
88981 ttagaactga aactgtagaa cgtgaaactg cagataaaga acaactgtgt ttgggattta
89041 cacttattag tatttgtagc tctagtgtat tcatttaac caattcatag catcatactg
89101 caaagtagat tacaatttat ttaccattcc cctactggac aacatgtaca ttatttctaa
89161 tattccatta aaataatgct gcaatgaaaa ccatttttc atgtctcctt atttatattt
89221 acagagttac agatggtcta tattcagctt cattatattt ttgtccagtt gctctccttg
89281 gtaattggcg ctcccatcaa caatgcacga attcctattg ctgcacaact ttgtcaatac
89341 ttggcattgt gagatttttt aatgtctgaa aatatgaagg gtaagagcat ccccctattc
89401 ctttatttta tgtttttgag acatggtctc gctctgtcac ccaggctgga atgcagtggc
89461 gctattatgg ctcactgcag ccttgacctc ccaggctcaa gtagctggga ctacaggcat
89521 gtgctaccat gcctggctaa ttttgtgtat ttttttgtag aaacaggatt ttgccatgtt
89581 gcccaccccc gatcttgaac tcatgggcct cggcctccca aagtgctgca atttcaggca
89641 tgtgccacca tgcctgctgc cactattcct ttaatttgca tctccccaat aattctgaat
89701 cagcatattt tttccagtgt taactggtta tttgggtttt ctttctatgg actccatatg
89761 tacatccttt gactattctc ctattgggtg gtctttttcc aattgtttgg aaagttctta
89821 catagtctag aattgaacct tttgtcagtt acacttaata cacattatct tcttcagttt
89881 tatcattttt tactttgttt atggcacctt ttgatgcatg gaaatttaaa atgtagtggc
89941 atttaatcaa tcttttcctt atgacctttt gtgtcttact aaagagatca ctatctacca
90001 caaggacaat caggattatg gcttctgtct ccattcacag agttttgttt ctattctgct
90061 tctctttcaa agatatactt agtttggtta aaaccaaaat gacacacagc tgtgactttg
90121 aaaatgttct ctcttgtatc acagctatgc atctgcaact gaaagtggtt tctaatgtgt
90181 aatttaattc tatctcaacc gaaatttcca catttcttta aatttctaaa attagggtaa
90241 agaataatcc acccaacatg cctatatttt ctattcctaa ttatctgatt cttggaccaa
90301 aagtcggctc tttctcatga gtctatgtta gcttcataaa caagcccctt caggtcaatc
90361 tttacttcat tatcctcctt tttcttcaca ttacttctca ttatttaaaa ttatattgac
90421 ttgcttattg tatatttctc tcatcattac atcatatgct atatgaaaac agatagcctg
90481 ctatatcccc aacatgtaaa ataatggaga ccaaggggca gagatataat ttgtctcaga
90541 tcttctcaac cttagttcaa tctcttccca tctcaaatat ttagtctctt aagcagtgtt
90601 gttttcctat gaattacaaa atataaaatc ataacaaaga tcactaacat tttaaaatga
90661 aaagcaaatt ctaataccaa aaaattaaac ctttaactta aattaaccaa aattaaatta
90721 agctttaaag cagccacaaa ttgctttgtg ccatcactag aagaaatatc tctgctattc
90781 agaggagtaa gggcccccat cacattttgg ccctaactaa aacaaataca gtaggaaaat
90841 gcaccaaaag gaaaacaaat tatgaggtag ctttgttctt ttaatgggtt atatacttta
90901 aaaaatagtc tctatttttct aaacttcttg taatatggtt ttataactta aaattttaaa
90961 ttttttcatt caaaaacaga aaaaaaacca aaaactagaa aacaaggata tttactgacc
91021 ctttttcaac attttctttt ttgttgtagg aacagatgaa gttaatgcac atgcaaatgc
91081 cacatcttct ataaccttag aagatcctgc tgataaaaca acaaagaatc ttaaattttt

Figure 28 Cont'd

```
91141 ttcaaattat ctcctgtgct agagtacagg tactttctaa gggtaagaaa tagaatatca
91201 ttttcaccct ctattgtaag agcttatatt attaagttaa aatatttggt aaccaaaatt
91261 actacctctg acactaagga tatcttttca actcagagaa acgatgtagt cattcatggt
91321 aagacttgag gagaagaaat gaaagtagga gcagtcatat gcaaagtaat tcggctctca
91381 ggtggaagga gaatttttca aggaaacaaa cagtttgatt cctaatgata aggctgccaa
91441 attttcagat cagtcaagat gatcagttat tttccaaatg caatacaatt tctcaacgaa
91501 tgagaaatat agtatgatgt gaattacagg aaaataaaac atattttaaa gaaaaagaaa
91561 aaatttgatt attttcacat ctgtataata atgtaaacaa taatagtatt gccaggaaga
91621 cattctcaac ctaattttgt aattactatt catactacac acataggttt tattgaagtc
91681 attttatagg tagtagaaaa ttgttttaag tatagtttaa atattattat cccttagatt
91741 ttaatttgca actatttatt cacttttcct cctttacaaa attaatttaa aacccatggt
91801 tagctccatg aagagttcaa atgtaaacat cagccaagtt taagagctat tactaaaaca
91861 ttccaaccta gcgagataca aatcaatgga gctgtacttt aataatcttg aagtctagat
91921 tatgggtata tagtatacac atacctttcg ccctgccttt agtttcagac tgtacatttt
91981 tcagtttgcc taaaaataaa acaaaaataa acataaaaat gggcaaaaca ggacattaga
92041 atataggaca cattatcaaa tcaaatctat tacatgttac ctttcatttt ctgcggcaat
92101 tcacttgttt caatttcctc tgaatcactg ctttctatgt actggacagc agtttttctt
92161 ctaaaattaa gtatacacaa agaaattttt ggaccaaccc aaagcaaaat aagatataca
92221 aacatgttta ataaatgcaa taaaaaactt ttgtcctact catttatttt ttgaaaaatt
92281 ttttaaagat tttgacaaat gaaagtcaga tctgctaaaa gcattctgct ttcataacca
92341 gatagatgca ttaactatac aaaatggagt accaccgtta ttgtttttg agacagggtc
92401 tcactctttt gcccaggatg gagtacagtg acacaatcag aagtcactgc aacctcatac
92461 tcctgggctc aagtgatcat cccgccttag cctcgggagt acctaggact acaggtgcgt
92521 gccactacac ccagcgaatt tcactattat taaaataaat ttgatagccg ggcgtggtgg
92581 tgggtgcctg tagtcccagc tactcgggag gctgaggcag gagaatgacg tgaagccagg
92641 agacggagtt tccagtgagc ggagattgcg ccactgcaca ccagcccgcc tgacagagcg
92701 acactccatc tcaaaaaaat aaaataaaat aaaataaaat aaaataaatt tgaaaaacta
92761 agtcattcta aaaaattaaa aatttcaata ccaggtttaa tatttggtag ttcaattaaa
92821 gcatataaaa ctggaacaaa aaagaattgg ctcgtggcca ggcacggtgg ctcatgtctg
92881 caatcccagc actttgggag gcggaggtgg gcagatcacc tgaggtcagg agtttgagac
92941 cagactggct aacatggtga aaccccgttt ctactaaaaa tacaaaaact agccaggtgt
93001 ggtggtgcgt gcctgtagtt ctagctgctc aggaggctga ggcaggagaa ttgcttgaac
93061 ccgggaggca gaggttgcag tgagccgaga tcatgccatt gcactccagc ctgggcgaca
93121 ggagcaaaac tctgtctcaa aaaaaaaaaa aaagaaaaga aaagaaaaaa gaattggctc
93181 ttatcaaaga aaatacttct aacacaatga aataagcagc caaaaggcaa aaatgtaatt
93241 ctttcacaat aggaaatttc ataaattttg tgtagaaagt ttaacagtga ttaagtgcct
93301 aagagcaata aattctgaag tcagactgcc tagctttaaa tcctcattcc accagtaatt
93361 tgttatgtga ttttaagtaa attacctaat attcttcaat gcctgtttcc tcataagaaa
93421 aatagaaata aaaatataat cctcatagaa taaaatgtta ctacatgcaa aagcccctaa
93481 aagagtgcct gttacacaat aacataacta ttagctataa catgtgtgtc agctgttgtt
93541 aagtttttct catgtcctca ccattattat ttctatgtcc tttaatctga tcttgtccaa
93601 ttaacagtgt tgtgttaccc aggtttatga taagttcaaa gttgtaacat ttgtcattta
93661 tctaaatctg aagcaattag atttagataa cactgatgag tgctggcttt gtgcttagtt
93721 cagacagcta ttaatgacag aacaattaac atagccagtt tcaaaacctt tgttgatgaa
93781 atagctagca gagacaaatt tgcgatgact tgtgtagata ctgatggtag acaatacaac
93841 aaaagcccaa tttttagacc tacttcagaa gataaatgaa gtttgtaaag ttaacagatt
93901 caagacaaga tcagagggga attacatctt ttattcaaat tctccctcta gtctatgtta
93961 atataatact gtccacaaat aactttatca tgtcttattt ccactacatg tgaagaactg
94021 aaacgatcaa gaatgctttt gtcaagttat ctaataaaat attcttcagt tttcatcccc
94081 ttaaatccat aaagaaaact tctaaaatct caaaaaaaag caaaaatccc aagatctaga
94141 tcttagttcc ttgacaaaat agaaaaaaag tagtagaaag atgagtgatt attaaaattt
94201 accaattcat ttcttttata ggtcagtatc attcttcttc atctgtaaaa tgagaatact
94261 ctatctgcct aagtcatagt gatcttggta gtttcaaata aaagaatgta tttgaaagaa
94321 ctacttaaca aatataaggg tcatggtcat ctctatctct ctctttttct ttttttttg
94381 agacggagtc tcactctgtc acccaggctg cagtgcagtg gtgcgatctt aactcactgc
94441 aagctctgcc tcccgggttc acgccattct cctgcctcag cctcctgggt aactgggact
94501 acaggcgccc acgaccacgc ctggctaatt tttgtctatt tttagtagag acggggtttc
94561 actgtgttag ccaggatggt ctcaatctcc tgacctcatg atccgcccgc ctcggcctcc
94621 caaagtgctg ggattacagg tgtgagccac cgtgcccagc ctctatctct tttacacata
94681 agaaaacaga gtctaaataa acatatgaca taatcaaacc ataatagttt acctttgggg
94741 gcgggagcta gacaattctg acatgcgaaa cttactcttc tccttgatat ctgaaatact
```

Figure 28 Cont'd

```
94801 gggttgttca ctacatctag atgcgtctat ttcaaagaaa aatgcaaata taaagtatta
94861 gtaaggtgtc ttagaaacta gtattatccc tcttaaaacc ttggttttct aattactgat
94921 tttcagtttt cttaagatct tcattccttt cattactcta ttttctctac ttttttccta
94981 atctagtatt tcctcatttc taattagaaa gtaccttcaa atttttatga cccagtcttt
95041 tctttgtact catccagccc atccccacac tagctgaatg tcttgctgta ccaaggcagc
95101 taagagctat tggaaaaaaa aatcaacaga accacaaata ggtgctgtta taaatttaca
95161 tgttcaaatc tcagatgtca ccttgtgtca cttggtaact ctttttacta aacactgttc
95221 ggctccccta cccatttatc agaatgatcc ataccagaat gatccatact acattgaccc
95281 ctttgcctag attataactc cttttattcc aagatgacct atttcacttt acaaagaaaa
95341 tgagaaggta tcttccataa aaatctcaat tttcctatct tttaatccac tcatttatct
95401 atactttcct tcaaccttac ttctttcctt ctactgctaa ctcatctacc taactcatct
95461 aggtagatga gttagaccac agaacaaaat caaaaacaaa aaaaacctca tcttacccta
95521 actcgccacc tttctatttc tttcctgttc caaacttaag taatagcacc aattgttatt
95581 tacaattttt cacctactac tcttaaaata ttttcagtct gctctaatct atctttagaa
95641 ataagcttta tgagggtcat ttaagtaaaa tccacagatc tctttccatt tctcctatga
95701 cttcagaatt tggctctaaa gctagctttc tagcattatc tgctgttatc catgcatatt
95761 acattttaaa tacaacaaaa ttcactgagg tcccagaata aaacatgctc tttttaaagt
95821 acccttttcat gagatttttt ttcacaaccc aaaatgctct tccccagttc tctaaatgtt
95881 atactaccca attcaatcat ctccttcatg aaacctttcc tatcagtact cagggaagac
95941 cttttaaaaa agagatccta aatttttgta cagtacaaag ccaaaactag aattacaata
96001 ttaggctgaa aaaacgcaaa tcaaatattg ggtcatattc aaattaaagt ttaccagatc
96061 aaatattgga tcatttcaaa attagcattt acctttgctt agtccatctg ccttttcact
96121 ggtattgttt cctccaagtt tcatagagtc atcgttaaca ttatattcct gggtaaatag
96181 gcatatttct taaacagtac tgctatcagt tttaaataac tttaaatgca tcctgtttta
96241 tattgcttta ctttgcataa taaaaacgac ttaaattgct acccaattct ggagcaactg
96301 aaggggaggg tgaataagtg aagaggaaaa tatacaaact gatttaaaat tttttcacgg
96361 gttatctcaa aaaaaaaaaa ataaataaaa ctacttggcc ttcacctttt cgaaaggaat
96421 accgaaagca aatctcaatt ataaatgaca cattattgtg agaaaaattc aagaactggg
96481 atttacttta taattaagaa ttctcagact tcaaactttt caaaccaatg ttccatgttc
96541 agatctttaa tgaggctgcc attgtcatcc tctctttctc tagaatttgc agggcaaata
96601 actctaatat ataatagtac aggtcccaga gtgtcctctc ttcttactcc atggggtatt
96661 ttcctgctag cccattttct tttcttttcc aatgtacctt tgtttccaaa ggcaccaatg
96721 tacatgaatt gaaaacaatt ttataaaatt tgaattgtaa gcacttaaaa aatccccccc
96781 tagagttaag tcttactaaa aatgtggttg aaaaggttca tcctctttaa aggtaccaat
96841 caaaaaaaac taaaaatctg ccgggcgcag tggctcatgc cagcactttg ggagaccgag
96901 gcgggcggat cacctaagat caggagctcg agactagcct ggccaacaca gcaaaaccct
96961 gtctctacta aatagacaca gattagccag gcatggtggc gggcacctgt aatcccagct
97021 acttgggagg ctgaggcagg agaattgctt gaactgagga ggtggaggtt gtagtgagcc
97081 aaggttgcgc cactgcactc cctctgtctc aaaaaataaa attaaattaa tttaaaatcc
97141 caaatacaat atttttagag atacgtaact agaagagtta ctttagacag ttaatgtcta
97201 agcattagct ccaagaaaga gatggtaaaa atgcttttca tcaaacatcc tatcaagtct
97261 acaataatta ctataaccat ttatattaaa cagtatgcaa aaaaactaaa tgtttattca
97321 aacataaagg atgtcctttg cattcatctt ctttttttct tttctgagac agagtttcac
97381 tcttgttgcc caggctggag tgcaacggca caatctcggc tcattgaaac ctctacctcc
97441 tgagttcaag tgattctcct gcctcagcct cccaagtaac tgagattaca ggtatgcgcc
97501 accacaccca gctaattttg tattttcagt agagacaggg cttcatcatg ttggtcagac
97561 tggtcttgaa ctcctaacct caggtgatct gcccgcctca gcctcccaaa gtgctgggat
97621 tacaggaatg agtcactatg cccagctttg cattattctt tatattaaaa tattttgctc
97681 agcaaacttt ttctctataa agttcaaaga taaacttttt atcaagtcca agctggatga
97741 aataatagta tgttagaatc taggggaatt taaatcaata tttcagtcta cttatgtttg
97801 gaatatctaa acatacttga aagagtgaca attttttttt ttttttaat aaatccctgg
97861 ctggaaatca catcttcagc tgccatttca attaatgtag caacccaaat attctaaaac
97921 ctttctacta taaaacacac tttaaatata taacttatct tggggaaaaa taagagaaat
97981 ccttattgtt aattgataat ttctcaatta ataattgaga ctaaagcagt aagtagtctt
98041 atgtgcttct atcaaatctg gtaacctaaa acttcagttt caataaccaa atgtaggaag
98101 aacaagacat aaaacattag gttgctcagg gtgaggaatt gaacctgaga cttcaacagg
98161 tttaactctc aataaaagta gaaagtagga agtaaaacta agtgcagaca ttaataaaat
98221 aaaatctcaa tagcactgac aggaaactga tgtactgcct tctctgaaaa ctaaaaacta
98281 ggaagaatgg taattcagaa gagttcagaa aagaagtata ttgcatacat ctttttgcta
98341 tttccttcta tttaagttac ttgaattttc aggtaagaaa agaagtaact gttcctctct
98401 agtaacattt atttttctca ctaaactaat ttttaattct tagaaaccaa atataaatgt
```

Figure 28 Cont'd

```
 98461 atgtttatat tatatacata tatatatgta tatatataca cacacacaca caaattttca
 98521 agaaaagtca cccttattct catcactagg acataagtta ttattaacat tttggtacag
 98581 atagtgtagt gaatggactt aaaacacaca tgaatatatt atttacaaag tgaggaacca
 98641 aactccacac aaagatttgt acctgtcatt ttttaactaa aaagaatcat gaaagtgatt
 98701 attcatcaaa aataatactt aatggttaca taatagtaag ttttatggct atatcgttta
 98761 actgttccct tgtcaaaaga tgtttcaaat tgtccacctt tattttattt ttatttgttt
 98821 tagagacagg gtctcacagt gtcactgaag gtggagtgca atagtgcaat catagctgac
 98881 cataggctca aactcctggg ctccagcaat cctcctgctt cagcctccca agtaggtgga
 98941 actacaggtg ccagtcacta tgcccagcct ttaaaacaat gttgagggaa aaaaaaataa
 99001 taataataaa aaaaaacctg tattttcttt ggaaaaaatg aatttctggt aactcctttc
 99061 agatagtttc ttagaaatgg cattattgac tggcaaaatt gttcaaagat tataaacaat
 99121 tttaagactc ctgatacaaa ttatcaattg ttctcaagaa tgtgatacct gctggtggct
 99181 cacacctata gccccagcac tttgggaggc tgaggcaggt agatggcttg agctcatgag
 99241 ttcgagacca gcctggacaa catggcgaaa tcccacgact tgagcctggg aagcagaggt
 99301 tacagtgagc tgagattgcg ccactgcact ccagcctggg caacagagcc acaccttgtc
 99361 tgaaaacgta aaaagaatgt gacacgaact ttcattctta tcagcaacac ataaaagaca
 99421 gagagaccat tcatcgcatt tctgaaaaca caatttcttg gtagtttttt ttaatgatga
 99481 aatttaaaga tttcacaata aactggaggt ttcacataaa gcgcacttac tactacttta
 99541 ccttcttcag tagattcttt ttaactcttt caataggaag aggtctgcca tcatggaagt
 99601 tggtaaggat tactgcaatg gccgtaagag ttttaccctt aaaaatgttt taaaaagata
 99661 aatggtcaga ttgtgaaacc cagttcacca gaaaaacgtt cccattttaa aatcaagtaa
 99721 tcgaatatca aatatcattc aaggtcacaa aagtgtactg tatattaatg aaaaatgagt
 99781 atttcattta gttagaaaat gctttccccc actgcattat tagggaatta ttaaatgatt
 99841 tacactataa aatactacaa gtatctctga gaagtgtcat tttattgaat gactgctatt
 99901 ctacaattct acatttaatc ctgaactctt gatttttatc atgcaactat ttataactac
 99961 tcatttgcca tgaagttcgg agctaaagtg attagtaatt ccagtatgca tataatcatt
100021 ttattaaaat ttgacttaaa gataatacca agcttaaaag tcataacttc tccataaagc
100081 ataaaagata tatttcttat gcatttttct taagtggaag tatgtaattt ttcaatattt
100141 actgcagctt aattcatagt tggcaataca ttagtaaatc tctggcaatt ttaaagcagt
100201 gtaaataaca gcagttttca aagtgtggtc cacaggccac tggggatggg gagggtttcc
100261 caggagccct tcacgaaaat gcaaagtcaa aactgttttt ataatactac taagacagac
100321 cttactctgt taacatctgc actaatccta caaaaagccc aaatagcaat ggtcaaaaaa
100381 gtgcctcagc ataaaacaaa gcagttggca ctaaggtata ttagtactag tcatcatatt
100441 cttaactgcc acatacagtt aaaaaaaaaa gttttcctta agaatgttct tggctgggca
100501 cagtggctca ggcctgtaat cccagcactt tgggaggccg aagcaggtgg atcacctgag
100561 gtaggagttc aagaccagcc tggaccaaca cggtgaaacc ccgtctctac taaatacaaa
100621 aaattagcta ggtgtggtgg tgcatgcctg taatcccagc tacttgggag gctaaggaag
100681 gagaatcact tgaacctggg aggcggaggg tgcagtgagc caagactgaa ccattactcc
100741 agcctgggca acaagagtga aactccatct caagaaaaaa aaaaaaaaag aagaagaatg
100801 ttcttgacga agcagtacac attaatttta ttaagtctca ttacttcagt acaaatcttt
100861 ttaatattct gtgtagtgaa atggaaacta catattaact acttctgttg catatcaaac
100921 tatgagggtt agctagagaa aaatcactca agatgggagt tataagctca actagtcact
100981 tttttcatgc gacaccattt ttacttgaat gacaactatc aatcacgact gagtgtctga
101041 aagatacttt ctagaaaatg aacaaagtga ggctgtcatc tacaaaggaa acaactgagc
101101 ttgtcaatga taaaatgtga gttttcaata gaaaattaga atgttggaaa agttgtgtgt
101161 gacactatga acatagtgcc ttcccaaaat gtaaatgatt ttctaataac atcagtggta
101221 atattaatga atgtaagttt tggatattgc atatgaaata catcaacatt tagtaggtct
101281 gcccaactca atgaaccaac attttctaaa tgagcaatgc ctggtgttat aaaatcatac
101341 atgggtaaaa tattcattca atgagcaaga tacacaaata gattttaaca aaacagggta
101401 taaaaagctc actgatatga tttagcgttc catattacaa caattaacct ttaataaact
101461 aatatttatt gaattttgat gcagtatatc aaagaaatct gcaattatct aaaaagctat
101521 tacaatactc ctctctttac agcaacatac cgtcgtaagg acaaattttc ttcatatact
101581 tcaaccaaaa caacctaaca gaacagactg aatgaagaag caggtgagaa ttcagctgcc
101641 atctattcgg ccaaacacag agatgtgcaa aaatgtgaca atgccagttg ttgcaacaaa
101701 aatgttaatg ttttggcata gttatttttc ttttaaaaat atgctattta tgttaacaag
101761 caatgagttt gtactgctat ctttagctac cactgcctaa tactataact actgataaca
101821 gcacacaaaa gcaaaagctc tttgaaatca gtaattttta tgagtagaaa gggttactga
101881 gactgaaaaa tataaaaact gctagtgtaa ggacacataa gtatctattt aatactctgc
101941 ataacacaat atttgtgatg gaatacaatt tcacataaca ttgttttgtc tttacttttt
102001 aaaagaaaac ctatgaaaga caggtcaccc tgaatttcaa agagcaaaga aaattgagaa
102061 acatggccag gcagggtggc tcacatctgt aaccccagca ctttgggagg ccaaggcggg
```

Figure 28 Cont'd

```
102121 cggatcactt gaagtcagga gttcaaaacc agcctggcca acgtggtaaa accctgtctc
102181 tactaaaaat acaaaaatta gctgggcatg gtggcacatg cctgtaatcc cagctatttg
102241 gaagtctgag acaggagaac, tgcttgaacc tgggaggcgg aggctgcagt gagccaagat
102301 cgcgtcactg cactccagcc tgggtgacag agcaagactc catctcaaaa aaaaaaaaaa
102361 gaaaactgag gaacacatct ttataccact attttaacta aaaactacag gaaaaactgt
102421 gttctacaaa gatagtaaga tctagtccca aactggtacc aacatattcc tatctctttt
102481 tacttcaaaa ttactttcta gaaaatcata atcacaaaat tagattttat taactaaaga
102541 aaaaaataat taccaaaccc atatcatcag ctaaaattcc tccatggaca ttttctggtc
102601 ggtccttctc agaaaaattt gttattgtgt tatagtataa gtcatttcgc tgttcccaga
102661 atggtggaag ttctttgcta ttttcccgtg acaccatcca agctagagct tgtttttgat
102721 gtggaagcag tggtgtttca atagcctata aataaaaagt cataaagcga aatacattaa
102781 gccattcctt tttcacatgc agatcctgtg attaaacaaa aagtagattc taagtatatt
102841 taaacattct gagtagtctt aaataaaaaa tattggttca atacctcagc tggttccatt
102901 tcatgggttt tatcatcttc ttttaaatct tcaaacaatt tgtcaaattc tgttttaagc
102961 tacaataaac agcaacaaga aacaatgtaa aacaaactaa ttaaaataat acttgcattt
103021 aagtcagatt tgaaatcatt taattaactg gcgttaatgt caaacgggcc aggagtggtg
103081 gctcatgcca ataatctcag cactttggga gaccaaggtg ggaggatcac ttgagctcag
103141 gagtttgaga ccagcctgga taacatagtg agagatccca tctctacaaa aaaattttta
103201 aaaaattagc caggcatggg ggtgcatgcc tgtagttctg gctcctaggg atgctgagat
103261 gggaggatca cttcagccta gtaggtagag gatagagtga accatgatta tgccactgca
103321 ctccagcctg ggcaagaaag tgagaccctg tcccaaaaaa ttaagaaatt gaaattaaat
103381 taaatttaaa aaccacaaac tagtcaaacg ccaaaaaggc acataaatca actctagaat
103441 tacactactc actgaagagc gaatgaacag gtgctttagt tgttaggagg taaaagaata
103501 taaaagaatt ttaagtcccg agtttcagtc acagttttat cacaaataag cttatatgta
103561 agtttcagca agacttgacc tttatggccc catgtcttat tcctataaaa tgaaagggat
103621 gaactaaatt tagattgtct ctaaggttcc ttctgccatt aaaattctgt gatgctatca
103681 gttgtattca aaatattttc tcagtcctac caggtcaatg acattccatt tattattact
103741 gcaataagga acttgccatg cactatagga tacttttatt attatgctaa cttaattgac
103801 tatttaaact aggttagtcc aattttatat caacatatat gtgatacaga tatttttgaa
103861 aattaacaca tgaacacagc agcagctcag aagaaatcag gccaggtcta aaaaatcaga
103921 agtaaccggg ggcttatcta ccagaaacaa gctcaagtcc ttctaacaat aagcagggta
103981 aatatggtaa ctacaattag tagataaaag gactaacact ggtagataca aacatcaaaa
104041 aagttttaat gttatttta aattgatctg cattaattac tttatatact tttttttcaa
104101 agaaattcag cataactcat cctaaattac aaattagaca atgtgaagaa tttggagtta
104161 aatctttagc actgcatttg caaaattact gaaaggagtt aagtctcagc ttcaattttt
104221 ttcttaaaat tactacttga acatttaaaa catttcagat acaaaaatat gccgtgcata
104281 aacaggaacg tgtgtgtttt gtatgttggg tacaaatgtt cacgtatgga tttattcctt
104341 atcagtactt tttatataaa atgttaacat atatatattc tatttattct ataacatgtt
104401 tcattataac ttaatatatt tgaagaccac aaatacccac acgtacatat tcatcctgtt
104461 cttttaaata caaacctgca acatttaaaa ttcacttta gattacaaaa taaaagagaa
104521 gcacacctgt tcagttgtca tctgtactgc agcatgcact ggcatactat agcttggtcc
104581 agctcttcca gagccccaac cactttccaa attgaatcct aaagctataa tttacaaaat
104641 aaaaagaata aagccatcaa ataaagtagg tttttatctt ttattaagta gcttttaaaa
104701 tatgcattaa cagggctggg tgctgtggct catgcctgta acccaagcac tttgggaggc
104761 caaggcaggt ggtttgcctg agtccagtag tttgagacta gcctgggcaa catggtgaga
104821 ttctgtctct acaaaaaata caaaaattag cagagcgtga tggtgtgcgc ctgtagtccc
104881 agctactcag gaggctcagg tggaaggact gctggagccc aggaggtcaa ggttgcactc
104941 agccactgat cgtggccact gtactctagc ctgggtgaca gagcaagact ctgtctcaaa
105001 atatatacac aaataaataa ataaaacgtg cataacaaac tgcattaaaa gccttgtatt
105061 ttagaggtat tttaaaagct aaatactaaa tcactaaaaa aacaagtaat tccacattat
105121 ttgaacaaac tcttgccaag caatcctttt cattcgctat cagaagatta tttttgccta
105181 aaatttatca aaaagaatgt ttaaagtcag gggtttccag atgttaccat tagggaaaat
105241 taggggaggg tatatatgag acttcattgc atatatactt tttctttaat ttttttaata
105301 tcccctcact cttactgtgt atatttttat gtgtgaatct atttcaatag aggcaaaagg
105361 aaaaatcaga gatttcctca aaaagactcc tgaaaaaaag ttgtaccttg gagccttgag
105421 cacaaattag ccaacaaact ccaagagaga ataaatgttt cataatccct gcatcttaca
105481 atatcagaat aatattaagt gaaaccctaa acttactttt tggtgcagga cccaatttaa
105541 atccatgttt cttcaactga tctgaaaccg cttttctatt ttcttctttt ccccaaaaag
105601 tcatatgcag aggcatggta aaagcattgt ttgcaccaaa aggaactacc ctattatatt
105661 tgggagaaaa agaaagggaa atcagaagca ttacttttta tccccactaa gaattagttc
105721 aatattttaa gtgtatcatt tacattttac attgaagtaa agtaataagg aactagttgg
```

Figure 28 Cont'd

105781 gcagacttga aagaaaaaaa ggaaaagaaa tagaaccata aacaatcaga agggtgattt
105841 ttagaaacag agtgccagaa aaaattgaga aaatggaata aagcagtttt aaaaatgcaa
105901 ggatggcctg taaacaagat tgtaatataa gtaaaataag ttaaatcaaa gggtagagtt
105961 tggctaacgg aaccactgag ggtagtaact cattctgctt ggaaaatggg aaatgcaaga
106021 acttgtgatt cactaagtta ttctcaataa ctgaaaaacc tgataaactg aaaaaggaga
106081 aaacaaagtg agacattatg agcatttatt ttctcttcaa actcacctag gattatatga
106141 atatcctaaa aattagaatt agaatatcaa gatcctaaat attccagata cttaaattga
106201 gatcacaagt acaaatgcct aaaagagaaa aggtaactga aagattgaaa gagtttagaa
106261 aaagaaaaaa gaaacaggtg ataaaactgg gagagcatat gccctagcta aagggagagt
106321 tccagtcaac tattgtcata tggaagtgac aactcatcct gcattgccag atttcccact
106381 tgtcaagaga aaaataagtt ttttgtattt aagaaaataa attaatatgt ataaaatgcg
106441 gacttagcta tttcaataat tagtactgaa cgaaaaaatg tctaaaacac tgcagggcaa
106501 aatgaaattc acctacatca tatcacagaa taccagttta taacctctga tataaacaaa
106561 cacttccagg ctgggcacgg tggctcacgc ctgtaatccc agcattttgg gaggccgagg
106621 cgggcagatc acctgaggtc aggagttcga gacagcctgg ccaacatggc aaatctctgt
106681 ctctactaaa aaatacaaaa attaattggg catggtggca ggcacctgta gctactcagg
106741 aggctaaggc agggagagtt gcttgaaccc tggaggcgga ggttgcagtc agccgagatc
106801 acacaactgc actgcagcct gggcaacaga gtgagacttc gtctcaaaat aaataaataa
106861 ataaataaac aaataaacac ttccaaattc ctaaacctag gtaccaaaag gtttggagaa
106921 atggaatgta cttcctaata tttgaaccaa ggtctatttt ctccatggta ataattccat
106981 aggaaaatga gaggacaatt gtttgatggc tggggtagta gaagtatgta aaaggagaaa
107041 taaaacattt atttaacacg taccaaaagt acagtgatag aaattgttag cataccttta
107101 taggctctaa caagaagccc taacaagttc aaagatgact ataattcaca attattaaaa
107161 ctaaattctg aaagtacatt acccttcaat ttgtgccaat ttgttgtcca tgatataggc
107221 caaagcacct gcaagctctt tctttaaatg gccaacttga tttccattca cattgtttac
107281 tttaattgca ttcttatcat aagggttatt aggatctcgt tgtaatgcaa ccatttcatt
107341 attattaacc taataaaaat gataaacaga taatattaca ttataaataa tagacaataa
107401 ctttgttata ctttatctgg gatttccatt catttttaa tgctactgct aaacactgca
107461 cataagaaaa atatattttc tagacacaaa ctacagaaat attaatatat attattatga
107521 ctccttaata ttatttttct aatctatatt cattaggaag ttaacaaatc cttaatgtct
107581 taaaaaatag tcaaatgtaa actagcaaat tcaaattatc taaaattttc aacatcataa
107641 cttctgaaaa atagccaacc cttattgagc gcttactatg tgccaagcat tggtctaagt
107701 acattgcaca gactgacttt ttaaatcctc acagctctat tgggaggtac tatcattatt
107761 cccatttttac aaatgaagaa actgaagcac agaaaagtaa cttgcccaag gtcacaaagc
107821 taaaaggtag cagagctagg atctgaaccc acacagggtg gttcctgagc ctcaattctc
107881 cagaatataa aagtcaggga ataaggcatt tataagtcag gatggaacta ggtaagaaaa
107941 atatatccag ccttcaactc atgccattct gattctactt tattaaaaac tgtatgattt
108001 aacttaaata ttatgatgta tatttttaaa aactgaagga gggtaacaca ataaagaatg
108061 gtagaatgag gactttcaaa aattcctcaa taaaggcaac aaaaaaacta gataaattat
108121 ttgtcagaat caaatgtttc agaacactga gaactaaatc aaaggcttgc agaaatctgg
108181 gtaacattta ttcaagaata aaaagaataa acagctgaat atcagtactt cagtcccatc
108241 cccagcaatc ctgtagcctt taaaaatagc tcacggccag gcacagtagc tcacgccccc
108301 agcactttgg gaggccgagg tgggaagatc acgaggtcaa gagatcaaga ccatcctagc
108361 caacatggtg agacccccgt ctctactaaa aatacaaaaa ttagctgggt gtggtggcac
108421 gcgcctgtag tcctagctac tcgggaggct gaggcaggag aaccacttga acccggcagg
108481 cagaggttgc agtgagccaa gatagtgccc ctgcactcca gcctggtgac agagcgagac
108541 ttcatctcaa aataaataaa taaataaata aataaataaa taaataaata aataaaaata
108601 atagcttgtt gggatcctag gtaaagcctg gcagtcacca gaaaaaaaga gaatggagtt
108661 acagttcttt cagagattca ttcccaaaga actgttattc tcctgaagtt ccccggaaga
108721 ccccacttgc aaggctgact gtatttaacc tctgagctca ccaagtacaa aaaaacctcc
108781 cttgggcgga tgtttgtcaa aacaatttta caggaaagtg ttttaacttc atggctacct
108841 gaggcagtgg ataacagctg aagcaaaaaa acaaaaaggc ttataaagaa gagctaggga
108901 atgagatgtc tgtgagggct ttgaaaagct ccagtgtatt tctggatatc tagaaggccg
108961 taagcaagca cagggctggt atgcatgacc agggctgtgc acattctcaa gaaagacctg
109021 agaaggccct aaacgctcac ctttgcctga acttgagcat ttaaacaagc cagaagtgaa
109081 agctaaagca gagttgtcag gggccttaga gtgttgaagg aatgccctaa catacagaag
109141 ttctcagcaa agaatgtacg atttattagt tccagcacaa tcatcagctg accgctaagc
109201 taaccaagta cagacttcag tgaccacaca cgataaagga tagacatcac agaattaatt
109261 caggaaagtc actaacaaac acacactaat tacaaaactc agcaacaaac caccctagaa
109321 gcctaacaaa cacacactaa ttacaaaact cagcaacaac aaaccaccct gatttccaga
109381 gctgccacat tatttaaaat gtcaattttt caagaaaaaa agtacaagac atgcaaaaaa

Figure 28 Cont'd

```
109441 aataagaaag tatggtctat acacagggaa aaaaagcaat caatgtcaac tgtccccaag
109501 aaagtataga tgtttgacat agtagaaaaa gaatttaagt cagttatttt aaatatgttc
109561 aaaggggctg ggcatggtgg ctcacacttg taatcccagc actttaggag gctgaagtgt
109621 gaggatctct tgaggccagg aatttgagac cagcctggga aacatagcaa gaccccattt
109681 ctacaaaata aaaatagaaa aattagccag gtatggtggt acatgcctgt agtcccagct
109741 attcaggagt ctgaggtggg aggactgctt gaatgcagga gttcaaggtt acagtgagct
109801 atgatcacgc cactgcactc cagcctgggt aatagagcaa gaacctatct ctaaaaaatt
109861 aaaaagttca aagacctggc cgggcacagt ggctcacgcc tgtaatccca gcactttggg
109921 aggccgaggt gggcggatca cgaggtcagg agatcgagac catcctggct aacacagtga
109981 aaccccgtct ctactaaaaa tacaaaaaat tagccgggcg aggtggcggg tgcctgtagt
110041 cccagttact cgggaagctg acacaggaga atgatgtgaa cctgggaggc agagcttgca
110101 gtgagccgag atcgtgccac tgcactccag cctgggcgac aaagcaagac tccgtctcaa
110161 aaaaaaaaaa aaataagttc aaagacctaa agcaaatcat ttctttcgac atgaaaacgc
110221 atgtctaaaa gcatggagac gatgtctcac caaacagcaa ttctagagtc aaaaggtaca
110281 ataactgaag aaaaaattca ccagagagac tcaacagcaa atttgagcag gcataagaaa
110341 aatcagcgaa gctgaagata ggtcaattga gattatacaa tctaaggaac aaaaaaatga
110401 cacacaacag agcctcagaa acttgtaata tcaagcatac caacacacat acataagagt
110461 accataagga gaaaagagaa aggggcagaa aaaatattta aagaaataac agccaaaacc
110521 ttcccaattc gatggaaacc actactctac acatataaga agctcaacaa ctccaactaa
110581 aataaactca aagagatcca cacctggtca catcataatc aaactgtcaa aagaagtttg
110641 agaccagcct gggcaaaagg gtaacgacag tttctaaaaa aaaaaaaaaa aaaaaaaaaa
110701 aaaaaaaaaa aaaattattt aaaaaaataa aaatttttaa gtataatttt ttgtaatttt
110761 tgtcctatct cattaaaaag tatataatta taaatctttg ttgataggta tactacatat
110821 ataaagatat aacttggata acagcacaaa ggaaggcaat aggatggagc tgtgagcaaa
110881 ggtttttttg gggttttctt tgttgttgtt tttttgagac agagtcttgc tgtcaccagg
110941 ctggagtaca gtcgcacaat ctcggctcac tgcaacctct gactccctag ttcatgcgat
111001 tctcctgcct cagcctccca agtagctggg actacaggca ggtgccacca tgcccagcta
111061 atttttgtat ttttagtaga gatggggttt caccatgttg gccaggatgg tctcaatctc
111121 ctaacctcgt gatccactca ccttggcctc ccaaagcgct gggattacag gcgtgagcca
111181 ccatgcccca tgcccagcca caaagttttt tatatactgt taagttggca ctaatccaaa
111241 ctaggttctt ataaattaag ttgttaattg taatccctaa agcaatcata agtaactcaa
111301 aatatatagt aaaagaaaca agggaattaa aatagtacac taaaaaaaaa aatctattta
111361 acataaaaat gagtattgga ggccagaccc aaatgccaca tattgtatca tttcatttac
111421 gtgaaatgtc aagaggaggt aaatccatat acacaggaag aagactgatt catggttgcc
111481 agggactgag gggtagagat aatgggagtg actacaaatg ggaatgaggt ttctttttgg
111541 agtaatgaaa atgttttgga actagattgt agggatggtt gcacaacctt gtgagtattc
111601 taaaaactac tcaactgtac actttaaaat ggtaaatttt ataataaatg tatctcaatt
111661 ttaaaaaatt ccttgaaaat atacaagaaa gctgcttcaa acctctccca gcatgtctac
111721 tgactcaaga gattacagcc agaactaaac agagtattag acagatattc atctctagac
111781 ttttagagat attatctatg aggcagagct atgttctctc ccagtcgcca ccctctgcc
111841 actgcctaca ggatgatctt ttttaaaccg gattatgcaa atcctagaaa atacataacc
111901 aaaataaaaa agcagggcac aaaactaaca atttttgagc atttcatctc aaggtattac
111961 ataattgtaa tttatcataa taagctatca ttaattacaa ttagttaact gtcgttacaa
112021 ataagtaatt actagaattt taaataaaat atatatgagc caaaatattg gagactataa
112081 aaacaaaaac atggtcccag cagcaagaaa aaacaggaca tataaaaaaa caaaaagaca
112141 aagacaagaa taacaaaagg aatgagaatt aatcataaca gtatagtcat accccgcata
112201 acaacatttc agcaatgaca gaccacatgc attatgcaac agttgttcca caagattata
112261 atggggctgc cctatagaag tgtaccattt ttcatatttt atacaatatt tttactgtac
112321 ttttttcatg ttatatatgc ttaaatacac aaatgtgtta caattgccca cagtattcag
112381 taaagtaaca tcttgttcac tctgtttcca caaagaaatc acctaacact tttctcagaa
112441 agtatcttca ttgttaagca acacataact gtattttaaa atagaaaaag aaataaacca
112501 cattataaaa taaagccttt gatgggtcaa aaaataaact taaccatatg cttattgaaa
112561 acacactaaa aaaaaaaaaa caaaaaataa aaatgaaata agaataagaa aacaccagag
112621 ttaaaaataa agaatgctac tacataacaa aacgtggtgg ctaaaggcaa cgcaatactt
112681 aacacgttta aagattaaaa gattaaatga ctcaaattaa gtgctcaaat cgcaagtcca
112741 acagaaatga aaagaaaata aatctaagta gaaaaaccga atgtaagagt aattattaag
112801 ccaagaaaca caatcattaa gaaaaattca agagagcctg ccttgaaaag aaataatgtc
112861 atagctaatc aataaaaaag acacgaatgt gtaacacaaa gaataaaaac aggaatataa
112921 tcaccttttt aaaaaagttt ttaaaaagtt taatggagcc aataaggcat ataaaattat
112981 gttctggcag acattaaaaa aatggcacat tttatcaaaa tttgacaatt cagagtaaag
113041 ggtactgaaa tttaaggttc agagcaccaa cttaaaacaa ccaaaaaaat ggttattcaa
```

Figure 28 Cont'd

```
113101 cttggaaaga gtaactagaa ccaaggcata aaagggagtt taagtatgtg ttcctccaat
113161 tcaaaatcct tcaggaatag gagttctggc cacttgtgag caacacatca taccaccaaa
113221 ccccaattcc agaaaattac agttacaatt ttataaggat aaatgattac tatgcacaat
113281 tctatgggaa taagtttgaa tatctaaata aaatggatgc ttttcttcaa aaatgtaact
113341 aaccaaaata gaccaaagaa gtacctaaat tatgctgagt gaaaaacgcc agacttaaaa
113401 tagtacatac tgtagaattc aatgtatatg aagttctaga ataggcaaaa cttatttata
113461 gtgatagaag acttgtgtgg ccgggcgcag tggctcacac ctgtaatccc agcactttgg
113521 gaggccgagg cgggcggatc acgaggtcag gagatacaga ccatcctggc taacacggtg
113581 aaaccccgtc tctaccaaaa atacaaaaaa attagccggg catggtggca ggcgcctgta
113641 gtcccagcta cttgggaggc tgaggcagga gaatggcatg aacttgggag gcggagcttg
113701 caatgagccg agatcccgcc actgcgctcc agcctgagaa acagagcgag actccgtctc
113761 aaaaaaaaca aaaaaaaaga aaagaaatga gactcgtggc tgcctagggc aggaacaggg
113821 aaaaatgaat gcaataggat acaagatact ttgagggaga cggatacaat ctgtatcttg
113881 attggggtag tagttatcag gtatacatct ttgtcaaaac actaggatca tacatttaaa
113941 atatgtacat tttatcatac gtaaaccaca tctcaataaa gtagaaaaaa aattaaaata
114001 acaagtatgc aaagatagta tctcaatttt tattattttt ataatagtga ttaactcaaa
114061 aatgtccatc aacaaggaaa ttataagtcc aatatagatg atattcctaa tcacattagg
114121 cataatatga tctcattttt attaatacat aatatgattc caatttacat gcacatgctc
114181 acagaaacaa gtgtagaaga taagttatgg ttctggatga tagacttcag aagatttaat
114241 gccttttag ccatttaaaa ataatgaacg tggctgggcg cggtggctca tgcctgtaat
114301 cccagcactt tggaaggccg aggtgagtgg atcacctgag gtcaggggtt caagaccagc
114361 ctgaccaaca tggcaaaacc tcatctctcc taaatacaaa aaattagctg ggcatgatgg
114421 cgcatgcctg taatcccagc tacttaggag gctgttgcag gagaaccgct tgaacctggg
114481 aggcggaggt tgcagtgagc caagattgca ccattgcact ccagcctggg caatgcaagc
114541 aagactccgt ttcaaaaaaa aaacaaaaac aaataaaaga acatgcatta attataaaat
114601 aaatgatatc tcaacatagt gttcctggaa aaaaataaaa tttaaagtat atatatttgg
114661 caaaataaga aaactgatga ttacttttta tcaacacaca ctagagttta ccgaagtcat
114721 agagattatt ctcagatcaa aaatactgat tgttctaaga tcaataacta ccataatcta
114781 gtatttcatg gtggtataaa ataattatcg aaagcaatgt aagttttgaa aaattagcac
114841 aaatctgaaa cagaaaagct cagaaatgca tatgagagcc agagaatata tagatgtgaa
114901 ataagaactg atagagttca agaatctatt agaaaaaatt agaggccggg tgcagcatcc
114961 cagcacttgg agaggccaag gcggacggat cacttgagcc caggagttcc agatcagcct
115021 ggacaacatg ccaaaaccta aaaaaaaata caaaaaaaaa aaatacaaaa attagctggg
115081 tatggtggtg aacgccagca gtcccagcta ctcaggaggc tgaattggga gaattgcttg
115141 agtctgggag gcagagattg cagtgaggcc agatcaggcc actgcactcc agcctgggca
115201 acaaaacgaa acttggtctc aaaaaaaaag aaaaaaaaaa aaatccaatt cacttaagtg
115261 aaagaaaatc acactggcct gagaattcta tgcaacatta ggatatggag attttattac
115321 aaagatttta ttacaaagat atattacaaa gattttaaat tcagcccaac tgtcattaag
115381 tccttcataa ctttcatttt atgaaactgt cattcataaa atgccacaag tatgaaaata
115441 taaaaactta ggaaatattg tctcttccta aggaatctcc tagacaacac atttcagaca
115501 atcaaaaaac aatttgaaac gttccaatgt aaaggaaata caggagaaaa acaaatccag
115561 aacatgaaac tcctaaggaa tctcctgtcg aacatgcttc agacacccta ttgagaagct
115621 taaacataag aaacatacgg agaaatgaaa gaataaaatc cagaagattc cacaaaacga
115681 tctccttttt gcaacaaatc gatgacatga acaaggcgct ggggtgtggg ggtggggtgc
115741 tacccataat aaccaaacat aacgcactga ggccaagcat attggctcat tcctgtaatc
115801 ccaacactct gggaggctga ggtgggaggg tcacttaagg tatttgttat tttttgcgta
115861 caaatattta actaacttaa agtaaaataa agattaaccg cacacatcac ctgcgaattc
115921 tcttattttc taggttaacg ccagaaatat aatcaaaatt taattatcta ctattatact
115981 cactactccc gtgtaatagc gtagtccaac cacatgacct ctcaaacttc caaataaaac
116041 ggaatctact tcttcatcac tagttagaaa gtcatctgga gggataacat cttggaattc
116101 aaaacgtgga aagaaagttg gatatgagag gcgtggaaaa tttccatgaa ctccatactg
116161 gacagtctgc aagtacttcc aaactggatc cctatttttt tttaaaggca aagaaaaaca
116221 atataatatt taagtatttc caaagaccat atgagtagtt ttctcatgct ttacttcagc
116281 aaatgaaatc ttacatggga atcacggtaa agattaatag ttactctcgt taaagcagga
116341 attgcagtga gtggagaagg cttctaggaa ctctacctgt tttcaatatt ggctcttcta
116401 cctgcccccc aggtattcaa aatgaaaagc ctagtcagag ttcactaacc tctttcccag
116461 tgaaaatcat ctaaattgat atgtatatga tgtctgtaac tacaattatg ggccggatgc
116521 ggtggctcac gcctggaatc ccagcacttt gggaggcaga ggcaggtgga tcacctgagg
116581 tcaggaattt gagaccaacc tggccaacat tgtgaaaccc tgtctactac aaatacaaaa
116641 attagccagg cgtggtggcg cccatatgta atcccagcta ctcgggaggc tgaggcagga
116701 gaatcacttg aacccgggag gcgtaggttg cagcgagccg agattgtgcc actgcactcc
```

Figure 28 Cont'd

116761 agcctggacg acagagcgag actccatgtc aaaaaaataa tttaaaaaat gaaatcatat
116821 gccaaaattt caatagggat tgagctatat aaaagactta taaaaataca actattaatg
116881 taccagaagt tgtttataat tttccagaaa atacatccag catccatcac tgtaactccc
116941 tgcgagcaga acacctattc tggtatccaa accacgccac cgtgggaaat tggccccaag
117001 ccacccgtac ccgccttccg tcgccggttt aagcataacc ccaaacccat taggtgtaac
117061 tgtttcggga atgaccataa ataccaattt ggacaatgaa aagtgaaggg gagtctactg
117121 gagcattgca aaaagtttcc attctcctaa gaacagtcgc ttatttgcga tgcgttttcc
117181 ttgcaagtgg cgcatatctc tactagcgcc tctgcaaata cttatctgca cgtctgtgca
117241 tacagataag ctgtgagcca cttggggaca aggactgtct tactcggccc tccaccgagc
117301 acaatgcctg gcgcacagtg ggtatctaat ggatgttttg ttaaataaag cactggccga
117361 gatgctttaa gccccaggcc ccccacagtc ggtgacagag atttcccaag tccctaacac
117421 gggactcgcc ctaggagccc ctactcgcca gcgaagacaa tgcatttatt tctccggcgg
117481 ccacatatgc gaccaacaga acgaatacag ctgcacaaat cgcccaggga acgcagagga
117541 acgcggggaa ggtcaggttc atttggggac gcctccaggc cgttagaccg agcgccccac
117601 cccctccgcc cccttcacct cttgaacatc caggacatgg cgctgagtgg gatgacaaga
117661 ggagcgcctc ggctcccctg gatcgttttc gagccgcctc gatacgcctc cttccaggcc
117721 ccgcagccct gaagccgggg acaaattccg agcgccggat caggagcgca cgactgaaag
117781 gtaagtcgcc gcgagtccag tcagacgtcg acgccgtctc cttctgcaac aatctgggag
117841 accagcgtcg ctctgtgact ggcactagga aagcccaatc acgaagagga gagtgcggag
117901 ccaaaccagt cagagcacag aagggagggc aactccgccc cgctgccatt caaagacggc
117961 gggggtccg ggctgcaagg gtggttccat ccgggttctt ccccgccccc aaggcgggcg
118021 cgcgggaaag ccacgaggcc ccaggagtgc gactgcggtg cctgcggtgc cggtgttttg
118081 tttgattccc tgcctcaaac ggagggaaac gaccttcctt tatcctacga gtcctaagac
118141 tgaaccccat tctaaaggct ctaccgtatc cttcctctta ttttcttctc atctaatgtg
118201 gcacataata gggccttata ctaaaggagc tccacggttt acgagaccta gagctaccgg
118261 cgagatactt tacttcattc cctgtggtga accgtgggga ctttcgctcc aaattttcat
118321 gttaagcctc agcgtatgca tgagacacaa cgagtttgga aaatcttaaa tggaacttag
118381 agtcccctcc ccacctcttt ttgttatttt taaggaaaat tttcctttct tggtgcagga
118441 aacccatcac atgtttatta cagctatggg ggcgtttgcc tgaaatggtg gacgggacca
118501 ttttccccgt gggcacttgg ctgctccagc caagaggggg aggcccttgt tttcctcaag
118561 gaactgcaga ggggcgctct gaggccctcc atggctctct ttccagagtc tgaggtgacc
118621 ggaaggagaa tgcggccctg ggaccgtcaa ccttggacca gctgcagccg acgcctggca
118681 gggctggtcg ctttgcgttg aggaggctgc tgtccccgaa gctggccttt taatcgcaca
118741 gggcaggaag ctggtggtgg cgcccagctg cacaggcggc accatgtaac tgccagataa
118801 tacttgcgcg tcacagagag gtccatgtta cacgcctgtc agcacaataa tattaggtgg
118861 tcagcttttc ttttttcttt tttttgtttt tttttctttt ttattgagat ggagtcttgc
118921 tctgtccccc aggctggagt gcaatggggg tgatcttggc tcactgcaac ctccgcctcc
118981 cggattcaag cgattctcct gcctcagcct ccggagtagc cgggattgca ggagcccgcc
119041 accacgccca gctaattttt gtatttttag tagagacggg gtttcaccat attggccagg
119101 ctggtcttga actcctgacc tcaggtgatc cgcccgcctc ggcctcccag agtgctggga
119161 ttacaggcat gagccaccgt gccaggccgg tagtcagctt ttcaagacac atttgttcat
119221 tatcgtaaat aaactgtagt gatctctaat catgaaccat ggatgagcaa tagaatttga
119281 aacaatgtat tatttcattt gaccaaagtt gatgaggaag ataaagacaa tggcatttca
119341 aattatttta attgttgtat gttcttcttg aagtgtgttg aggcaaatgg caatacagtt
119401 cagcttttag tatgccagat tttaaataaa ttcttggaaa atatgccaga aattgctcaa
119461 attgatgttt tgtaagagta agaaagtcat gctcactaga caaagaaaaa attccaaata
119521 tgagaacata gctctttcag acttaagact ggccaggcgc aatggctcac acctataatc
119581 ccagcacttt gggaggccga gttgggtgga tcacctgagg tcaggagttc gagaccagcc
119641 tggccaacat ggtgaaagcc tgactctact aagaaaatac aaaaattaag cagacgtggt
119701 ggcacgcacc tgtagtccta gctacttggg aggctgaggc aggagaatcg cttgaaccag
119761 ggagacatgg gttgcagtga gccgagatca agccactgca ctccagcctg agcaacagag
119821 cgagactcca tctcaaaaaa ataaaaagac ttaagatcta tgaataatga ctgtcccatg
119881 gttaaagaat gtgctttgaa tgaactaaaa tttgctattt aaaggaatgg tatggaaagg
119941 aaaggaatcc aaaatttctc catccagcca cttcccagtc acaaacacac ttctcatctg
120001 caccccccagc cacacacaca cacatgcccg cgcgcgcgca cacacacaca cacacacaca
120061 cacacacaga acctttatgc aaattaatca tgtcatgtca ctcccctgtt tgattcagtg
120121 agcctgaaat ccaagaatgg catatgtggc tcttcctccc acagtatgtt ttctctatgt
120181 tatcaatatt tcacatccca gaaccaggag taaaacattc tttcccttaa tcattctttg
120241 ttttatattt aaagatcaag tacaatttgt actagtttga ttaaaatgtt acagcaatta
120301 caatttcaaa actattatac taaataatgt tttctgaaaa attaactttt ttggtttttt
120361 cttgatttat tctgataaca gcatcacaag tagatatgaa aaatgaacac ttgtaactgg

Figure 28 Cont'd

```
120421 aaaatgaact gtagggtggc ttgtggggtt tggctggtga gtaagaagga aagtggcact
120481 aaaaggacgg tggggaagat aagggccagg ttacatagga acttaagagt ctccagtaaa
120541 atttgtgttt taactgcaat ggaaagccat tgaatgtttc gagcaggagg ataacgactt
120601 gatttaggct tttaaaaatg ctggcagctc tgtggagaat tacaggaaac aaggatagaa
120661 gcaactgata gaaaattatt gtgttcagat aagagatggt ggtggcttgg aaagggaagg
120721 tgatgaagcc aagagaacca aaatgttcac tgataaattt aggtaggaat ggtatggaaa
120781 ggaaaggaat ccaaaatttc tccagccagc cacttcccag tcacaaacac acttctcatc
120841 tgcaccccta gccacacaca cacatgcccg tgcacacaca cacacacaca cacacacaga
```

Figure 28 Cont'd